

OHPO OHPO

INTERNATIONAL APPLICATION PU	BLISI	HED (IN	DER THE PATENT COOPERATION	N TREATY (PC	T)
(51) International Patent Classification ⁵ :	•			1) International Publication Number:	WO 94/1	
C12N 15/31, C07K 13/00, A61K 39/1 // (C12N 15/31, C12R 1:21)	102	A1		3) International Publication Date:	9 June 1994 (09	
	CT/CAS			(81) Designated States: AU, BR, CA, F UA, US, European patent (AT, B	E, CH, DE, DK, I	Z, RU,
(22) International Filing Date: 23 November (30) Priority Data: 9224584.4 23 November 1992 (2)			_	GB, GR, IE, IT, LU, MC, NL, P. Published With international search report.	Γ, SE).	
 (71) Applicant (for all designated States except US): LABORATORIES LIMITED [CA/CA]; 1: enue West, Willowdale, Ontario M2R 3T4 (72) Inventors; and (75) Inventors/Applicants (for US only): CHONG, 32 Estoril Street, Richmond Hill, Ontario THOMAS, Wayne [AU/AU]; 31 Taylor F W.A. 6009 (AU). YANG, Yan-Ping [CN/C 120 Torresdale Avenue, Willowdale, Ont (CA). LOOSMORE, Sheena [CA/CA]; 70 Drive, Aurora, Ontario L4G 4R4 (CA). SI Charles [CA/CA]; 189 Mabley Crescent, Texture of the content of the content	755 Ste (CA). , Pele [6 LAC 0E Road, N CA]; Ar tario M Crawfo IA, Dwe	CA/CA 36 (CA 36 (CA Jedland: 51 1709 12R 3N 51d Ros 50, Yuar 6, Ontari	/-];), 7 e ı,			
L4J 2Z7 (CA). KLEIN, Michel [CA/CA]; It vard, Willowdale, Ontario M2P 1B9 (CA). (74) Agent: STEWART, Michael, I.; Sim & M. University Avenue, Suite 701, Toronto, On (CA). (54) Title: HAEMOPHILUS OUTER MEMBRAN.	McBurn ntario M	ey, 33 ISG 1R	0			
Purified and isolated nucleic acid from specific strains of Haemophilus influenzae is provided which encodes at least a portion of the D15 outer membrane protein of Haemophilus. The nucleic acid is used to produce peptides, polypeptides and proteins free of contaminant associated with Haemophilus for purposes of diagnosis and medical treatment. Furthermore, the nucleic acid may be used in the diagnosis of Haemophilus infection. Antisera obtained following immunization with the nucleic acid D15 outer membrane protein or peptides also may be used for the purpose of diagnosis and medical treatment.		ALICOLONIC GNOTECACTE	FRE	DIS SEUENCE COMMISSION PENNICHMONOCOLECODRASIA-RACCE/TORNASIA-VASCASSISTOMONOCOLECODRASIA-RACCE/TORNASIA-VASCASSISTOMONOCASSI DIA DEPENNICHMONOCOLECOMONOC	ELEPLISALMINOTERSONO	Ca Engern HurmA SB13 PMK Ca Engern HurmA SB13 PMK Ca Engern HurmA SB23 PMK Ca Engern HurmA SB23 PMK
	M788	TASSWILD	PAI	PROTECTION OF THE PROT	TPINOMASIATAVITIKSDINY	Ca Eagen HirriA SB13 PAK Ca Eagen KirriA SB33 PAK Ca Eagen

PTHYSISONWISTYONOMISMSING TEDYNOCHOGOSIGNSTOCEOOSIGHURSONOMICUAL PRINCES.

Ca Eagun MumA SB33 FPK

30

35

TITLE OF INVENTION HAEMOPHILUS OUTER MEMBRANE PROTEIN

FIELD OF INVENTION

5 The present invention is related to the field of molecular genetics and is particularly concerned with the cloning of an outer membrane protein D15 of Haemophilus.

BACKGROUND OF THE INVENTION

Haemophilus influenzae type b (Hib) is a major cause 10 of bacterial meningitis in children under the age of five years. Protective antibodies to the disease are induced by the capsular polysaccharide of the organism and a vaccine was developed that utilises the purified polyribosyl ribitol phosphate (PRP) as the antigen. This vaccine provides 90% protection in adults and in children 15 over 24 months of age, but was ineffective in children under 24 months Zangwill et al 1993 (The references are identified in a list of reference at the end of this disclosure). Like other polysaccharide antigens, PRP 20 does not induce the proliferation of T-helper cells, and re-immunisation fails to elicit either a booster response or an increase in memory cells. Conjugation of the PRP polysaccharide with protein carriers confers T-cell dependent characteristics to the vaccine substantially enhances the immunologic response to the PRP antigen. Currently, there are four PRP-carrier conjugate vaccines available. These are vaccines based upon <u>H. influenzae</u> type b capsular polysaccharide conjugated to diphtheria toxoid, tetanus toxoid, or Neisseria meningitidis outer membrane protein (reviewed in Zangwill et al 1993).

However, the current <u>Haemophilus</u> conjugate vaccines only protect against meningitis caused by Haemophilus influenzae type b. They do not protect against other invasive typeable strains (types a and c) and, more importantly, against non-typeable (NTHi) strains which are a common cause of postpartum and neonatal sepsis,

10

15

20

25

30

35

otitis media, epiglottitis, pneumonia, and tracheobronchitis, are required.

SUMMARY OF THE INVENTION

The present invention is directed towards the provision of purified and isolated nucleic acid molecules comprising at least a portion coding for a D15 outer membrane protein of a species of Haemophilus. The nucleic acid molecules comprising at least a portion coding for D15 outer membrane protein are useful for the specific detection of strains of Haemophilus, and for diagnosis of infection by Haemophilus. The purified and isolated nucleic acid molecules, such as DNA comprising at least a portion coding for D15 outer membrane protein, are also useful for expression of the D15 gene by recombinant DNA means for providing, in an economical manner, purified and isolated D15 outer membrane protein.

The D15 outer membrane protein or fragments thereof or analogs thereof are useful immunogenic compositions for the preparation of vaccines against diseases caused by Haemophilus, the diagnosis of infection by Haemophilus and as tools for the generation of immunological reagents. Mono- or polyclonal antisera (antibodies) raised against the D15 outer membrane protein produced in accordance with aspects of the present invention are useful for the diagnosis of infection by Haemophilus, specific detection of Haemophilus (in, for example, in vitro and in vivo assays) and for the treatment of diseases caused by infection by Haemophilus.

Peptides corresponding to portions of the D15 outer membrane protein or analogs thereof are useful immunogenic compositions for the preparation of vaccines against disease caused by Haemophilus, the diagnosis of infection by Haemophilus and as tools for the generation of immunological reagents. Mono- or polyclonal antisera raised against these peptides, produced in accordance with aspects of the present invention, are useful for the

15

20

25

comprises at least an 18 bp fragment selected from the DNA molecules as recited above is inserted. The recombinant plasmid may be plasmid DS-712-2-1 having ATCC accession number 75604, deposited November 4, 1993 and plasmid JB-1042-5-1 having ATCC accession number 75006, deposited November 4, 1993.

The plasmids may be adapted for expression of the encoded D15 outer membrane protein in a host cell, which heterologous or homologous a incorporation into a recombinant vector, provided in accordance with a further aspect of the invention. recombinant vector may comprise at least a DNA segment comprising at least an 18 bp fragment selected from the DNA molecules as recited above and expression means operatively coupled to the DNA segment for expression of the gene product encoded thereby in the host cell. plasmid for expression of the encoded D15 outer membrane protein may be plasmid DS-880-1-2 having ATCC accession number 75605, deposited November 4, 1993 being adapted for expression at the D15 outer membrane protein in E. coli. The selected DNA segment may encode a polypeptide of at least 6 residues and, in particular, may be selected from those segments encoding a polypeptide of Table 2 (below). The DNA segment may further comprise a nucleic acid sequence encoding a leader sequence for export of the gene product from the host. The host for expression may be selected from, for example, Escherichia Bacillus, Haemophilus, fungi, yeast baculovirus expression system may be used.

Additional aspects of the invention include the protein encoded by the DNA molecule comprising at least a portion coding for the D15 outer membrane protein, fragment or a functional analog of such protein, the use of the protein or analog in vaccination and diagnosis, and the generation of immunological reagents. The invention also includes antisera (antibodies) raised

10

20

25

30

of the immunogenic composition or the nucleic acid molecule as recited above to provide protective immunity against <u>Haemophilus</u> infection.

The present invention further includes a chimeric molecule comprising a D15 protein corresponding thereto as provided herein linked to another polypeptide or protein or a polysaccharide. The linked polypeptide or protein may comprise a surface protein or peptide corresponding thereto pathogenic bacteria, which may be the P1, P2 or P6 outer membrane protein of H. influenzae. The linked polysaccharide preferably comprise a PRP molecule from H. influenzae.

BRIEF DESCRIPTION OF THE FIGURES

The present invention will be further understood from the following description with reference to the drawings, in which:

Figure 1A shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> type b Ca strain (SEQ ID NO: 1) and its deduced amino acid sequence (SEQ ID NO: 2);

Figure 1B shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> type b Eagan strain (SEQ ID NO. 3) and its deduced amino acid sequence (SEQ ID NO: 4);

Figure 1C shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> type b MinnA strain (SEQ ID NO. 5) and its deduced amino acid sequence (SEQ ID NO: 6);

Figure 1D shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> non-typeable SB33 (SEQ ID NO. 7) and its deduced amino acid sequence (SEQ ID NO: 8);

Figure 1E shows the nucleotide sequence of the D15 gene from <u>H. influenzae</u> non-typeable PAK 12085 (SEQ ID NO. 9) and its deduced amino acid sequence (SEQ ID NO: 10);

Figure 1F shows an alignment of the nucleotide sequences of the D15 genes (SEQ ID NOS: 1, 3, 5, 7 and 9)

15

20

30

35

97kDa); 2, GST standard; 3, GST-(D15 fragment) fusion protein; 4, fusion protein cleaved by thrombin; 5, N-terminal rD15 fragment; 6, GST; 7, low molecular weight markers;

Figure 10 shows guinea pig IgG antibody response to N-terminal rD15 fragment. The arrows indicate the immunization schedule. Bleeds were taken at 2, 4, 6 and 8 weeks. The bars represent the standard deviation; and

Figure 11 shows the hydrophilicity plot of D15 10 established by using a window average across 7 residues according to Hope, 1986.

GENERAL DESCRIPTION OF THE INVENTION

Any <u>Haemophilus</u> strains that have D15 genes may be conveniently used to provide the purified and isolated nucleic acid molecules (which may be in the form of DNA molecules), comprising at least a portion coding for a D15 outer membrane protein as typified by embodiments of the present invention. Such strains are generally available from clinical sources and from bacterial culture collections, such as the American Type Culture Collection. <u>H. influenzae</u> strains may include types a, b and c strains, non-typeable strains and other bacteria that produce a D15 protein, fragment or analog thereof. Appropriate strains of <u>Haemophilus</u> include:-

25 <u>H. influenzae</u> type b strain Ca;

H. influenzae type b strain MinnA;

H. influenzae type b strain Egan;

H. influenzae non-typeable b strain SB33; or

H. influenzae non-typeable b strain PAK 12085.

In this application, the term D15 outer membrane protein is used to define a family of D15 proteins which includes those having naturally occurring variations in their amino acid sequences as found in various strains of, for example, <u>Haemophilus</u>. The purified and isolated DNA molecules comprising at least a portion coding for D15 outer membrane protein of the present invention also

10

described herein are advantageous as diagnostic reagents, antigens for the production of <u>Haemophilus</u>-specific antisera, for vaccination against the diseases caused by species of <u>Haemophilus</u> and for detecting infection by <u>Haemophilus</u>.

Reference will now be made in detail to the presently preferred embodiments of the invention, which together with the following Examples, serve to explain the principle of the invention. For clarity of disclosure, and not by way of limitation, the detailed description of the invention is divided into the following sections:

(i) The DNA sequences coding for the outer membrane protein D15 from <u>H. influenzae</u> type b Ca strain.

15 A clone producing the outer membrane protein designated D15 of <u>H. influenzae</u> type b (Hib) was isolated by screening a genomic library with H. influenzae type b OMP-specific polyclonal antibodies as previously described by Berns and Thomas 1965; Thomas and Rossi 20 1986. The DNA fragment encoding the D15 protein was isolated, subcloned into pUC19 to produce pUC19/D15 (Figure 2) and used to transform E. coli HB101 as described in Example 1. Plasmid DNA was prepared from two individual colonies of E. coli HB101 containing the 25 pUC19/D15 plasmid. Sequencing was performed on an ABI DNA sequencer model 370A using dye-terminator chemistry and oligonucleotide primers which had been synthesized on an DNA synthesizer model 380B, and purified chromatography. Nucleotide sequence analysis of the D15 gene revealed that it contains a putative promoter and an 30 open reading frame encoding 789 amino acids (Figure 1A).

The first 19 amino acid residues of the translated open reading frame form a typical leader sequence as found in other <u>H. influenzae</u> type b outer membrane proteins, such as P1 and P2. The N-terminal sequence of immuno-affinity purified native D15 antigen was

15

20

heterologous proteins in E. coli. The T7 expression system is described in U.S. Patent 4,952,496. were, therefore, constructed which utilize the T7 system to express a mature D15 protein that contains additional methionine residue at the amino terminus. The D15 signal sequence was removed during this construction process. A full length recombinant D15 (termed rD15) was expressed in inclusion bodies which allow the D15 protein to be readily purified. The D15 genes from H. influenzae type b strain Ca and H. influenzae non-typeable SB33 strain have been expressed at high levels in E. coli using the T7 system to permit production of large quantities of rD15 protein. The construction of clone DS-880-1-2 which expresses the SB33 D15 gene is described herein (see Figure 4 and Example 5). The rD15 protein was immunologically similar to its native counterpart isolated from <u>H. influenzae</u> typeable and non-typeable strains (see below). Thus, rD15 may be used as a crossreactive antigen in a diagnostic kit to detect many, if not all, strains of <u>H. influenzae</u> and other bacteria that produce a D15 outer membrane protein or analog thereof. Alternatively, rD15 can be used as an antigen to specifically detect the presence of H. influenzae in a sample.

25 A truncated D15 fragment was expressed in E. coli as a fusion protein with glutathione S-transferase (GST), as described in Example 6. The construction was designed to express the N-terminal fragment of the D15 protein. fusion protein was expressed at high levels from a pGEX-2T construction and the N-terminal fragment was cleaved 30 from the GST carrier protein by treatment with thrombin. This procedure generated a molecule termed the N-terminal rD15 fragment which encompasses amino acids 63-223 of the This N-terminal rD15 fragment was highly D15 protein. 35 immunogenic and elicited protective antibodies against challenge with live H. influenzae.

15

20

25

35

urea (see Example 8). After dialysis against PBS to remove urea, more than 80% of the D15 protein remained soluble. This soluble rD15 antigen was used for the immunogenicity studies described below. From shake-flask experiments, it was estimated that about 10 mg of soluble rD15 protein was obtained from 1 L of E. coli bacterial culture. It is clear that growing the recombinant E. coli strains under optimised fermentation conditions significantly increase the level of rD15 production.

10 (vi) Immunogenicity of the full-length recombinant D15 protien (rD15).

The immunogenicity of the full-length rD15 protein was studied in guinea pigs and mice. Using the immunization protocols described in Figure 7, a 15 μ g dose of rD15 induced high IgG titers in guinea pigs when administered in the presence of either Freund's adjuvant or AlPO₄. In the mouse dose-response study, the protein appeared to be immunogenic at a dose as low as 5 μ g in either Freund's adjuvant (Figure 8A) or AlPO₄ (Figure 8B).

The protective ability of rD15 against <u>H. influenzae</u> type b infection was examined in the infant rat model of bacteremia essentially as described by Loeb (1987). Thus, infant rats passively immunized with guinea pig anti-rD15 antisera were significantly less bacteremic than controls injected with pre-bleed sera, which is consistent with the previous report by Thomas et al. (1990).

(vii) Purification and characterization of the N-terminal rD15 fragment.

The truncated rD15 fragment corresponding to the N-terminus of the D15 protein (residues 22 to 223) as described in Example 6, was expressed in <u>E. coli</u> as a soluble protein fused to GST. The fusion protein (46 kDa) was readily extracted using phosphate buffered saline (PBS). Purification of the GST-D15 fragment fusion

15

20

25

30

35

the immunization protocols described in Figure 10, a 10 μ g dose of N-terminal rD15 fragment induced a good booster response in guinea pigs with almost all the adjuvants tested. The highest anti-D15 IgG titer was observed in the group of guinea pigs immunized with N-terminal rD15 fragment in Freund's adjuvant. The second best adjuvant was Titermax (CytRx Inc.). The other two adjuvants, TPAD4 (tripalmityl-Cys-Ser-Glu₄) and AlPO₄ were equally potent.

10 (ix) Protective ability of the N-terminal rD15 fragment against <u>H. influenzae</u> type b challenge.

An <u>in vivo</u> challenge model for a assessing the protective abilities of antigen against diseases caused by <u>Haemophilus</u> is the infant rat model of bacteremia as described by Loeb 1987. The protective ability of the N-terminal rD15 fragment against <u>H. influenzae</u> type b challenge was examined in this rat model. As illustrated in Table 1, infant rats passively immunized with rabbit anti-N-terminal rD15 fragment antisera showed significantly lower bacteremia compared to those injected with pre-bleed sera.

Since passively transferred antisera against the N-terminal rD15 fragment were found to be protective in the infant rat model of bacteremia, it was of interest to identify the protective epitope(s) of this N-terminal rD15 fragment. The first nine overlapping peptides of the D15 protein as listed in Table 2 were chemically synthesized based upon the amino acid sequence derived from the sequence of the D15 gene from H. influenzae type b Ca (Figure 1). These synthetic peptides were assessed for their reactivities with either rabbit or guinea pig antisera raised against purified N-terminal rD15 fragment by ELISAs. As shown in Table 3, both guinea pig and rabbit antisera reacted with a cluster of D15 peptides, including peptides D15-P4 to D15-P8 encompassing residues 93 to 209 of the D15 primary sequence.

10

15

20

25

30

35

was completely blocked by the addition of this mixture of five peptides (Table 5, group #2, 106%, $p = 0.53 \times 10^{-8}$). These results strongly indicate that a cocktail of D15 synthetic peptides may be used as immunogens to induce protective antibodies against <u>H. influenzae</u>.

(x) Epitope prediction and peptide synthesis.

To map the immunodominant T-cell or B-cell epitopes of D15, overlapping synthetic peptides covering the entire D15 protein sequence (Table 2 - SEQ ID NO: 14 to 49) were synthesized using the t-Boc solid-phase peptide synthesis as described in Example 15. The peptides were chosen based on their high index of hydrophillic β -turns estimated by secondary structure prediction analysis (Figure 11). Such peptides are likely to be surface-exposed and antigenic. Peptides more than 25 residues in length were selected to better mimic native epitopes.

(xi) Identification and characterization of immunodominant epitopes of D15 using synthetic peptides.

linear B-cell map the epitopes of overlapping synthetic peptides representing the entire sequence of D15 were individually coated onto ELISA plates and probed with several anti-rD15 antisera as described in Example 19. The results are summarized in Table 6. Mouse antisera raised against rD15 reacted with all D15 peptides, but the major epitopes were located within peptides D15-P8 (residues 180-209 - SEQ ID NO: 21), D15-P10 (residues 219-249 - SEQ ID NO: 23), D15-P11 (residues 241-270 - SEQ ID NO: 24), and D15-P26 (residues 554-582 - SEQ ID NO: 39), respectively. Rabbit anti-D15 antisera recognized only peptides D15-P4 (residues 93-122 - SEQ ID NO: 17), D15-P14 (residues 304-333 - SEQ ID NO: 27) and D15-P36 (residues 769-798 - SEQ ID NO: 49). Guinea pig antisera raised against rD15 reacted with peptides D15-P2 (residues 45-72 - SEQ ID NO: 15), D15-P4 (residues 93-122 - SEQ ID NO: 17), D15-P6 (residues 135-164 - SEQ ID NO: 19), D15-P8 (residues 180-209 - SEQ ID

10

15

20

25

30

35

cells in the immune system is to provide helper activity for eliciting high levels of antigen-specific antibodies following immunization. Antigens containing Th1 epitope(s) stimulate antigen-specific T-cells to produce high levels of IL-2 and IFN- γ , whereas Th2 epitope(s) induce high levels of IL-4 expression. Th0 epitope(s) stimulate the synthesis of IFN- γ and IL-4.

Little is known about the cellular immune response to outer membrane proteins of H. influenzae and its role in the protection against H. influenzae infection and diseases. To this end, the inventors performed studies of the cellular response elicited in mice following rD15 immunization. D15-specific T-cell epitopes determined using D15 peptides and T-cell lines obtained from five BALB/c mice immunized with rD15 (see Example 23). The lymphocyte proliferative responses of the D15specific T-cell lines to overlapping D15 peptides were determined in conventional cytokine assays as described in Example 24. The results summarized in Table 7. revealed that stimulation only with certain synthetic peptides elicited proliferative responses and the release of specific cytokines. Synthetic peptides corresponding to residues 114-143 (D15-P5 - SEQ ID NO: 18), 282-312 (D15-P13 - SEQ ID NO: 26) and 577-602 (D15-P27 - SEQ ID NO: 40), and 219-249 (D15-P10 - SEQ ID NO: 23), 262-291 (D15-P12 - SEQ ID NO: 25), 390-416 (D15-P18 - SEQ ID NO: 31), 410-435 (D15-P19 - SEQ ID NO: 32) 554-582 (D15-P26 -SEQ ID NO: 39), 596-625 (D15-P28 - SEQ ID NO: 41), 725-750 (D15-P34 - SEQ ID NO: 47) and 745-771 (D15-P35 - SEQ ID NO: 48) were shown to be highly stimulatory for rD15specific BALB/c Th0 cells and Th1 cells, respectively. Therefore, these immunodominant T-cell epitopes can be used as autologous carriers for PRP, and/or OMP B-cell epitopes to enhance their immunogenicity. The Th1 cell epitopes identified above may be useful in the H.

10

15

20

25

30

linked to polysaccharides including PRP as synthetic glycopeptide or lipoglycopeptide conjugates to produce alternate vaccines. These vaccines can be used to immunize against diseases caused by H. influenzae when administered to mammals, for example, by the intramuscular or parenteral route, or when delivered using microparticles, capsules, liposomes and targeting molecules, such as toxins or fragments thereof, and antibodies, to cells of the immune system or mucosal surfaces.

(xiv) Utility of D15 as carrier protein for the production of glycoconjugates.

To determine whether D15 may serve both as a protective antigen and a carrier, D15-PRP conjugation experiments were performed as described in Example 14. The D15-PRP conjugates were found to be highly immunogenic in rabbits and able to elicit both anti-D15 and anti-PRP IgG antibody responses as judged by D15-specific ELISA and PRP-BSA immunoassay (Table 9). These results clearly demonstrate the practical utility of D15 as a carrier protein for glycoconjugation technology.

In preferred embodiments of the present invention, the carrier function of D15 can be generally utilized to prepare chimeric molecules and conjugate vaccines against pathogenic bacteria, including encapsulated bacteria. Thus, the glycoconjugates of the present inventions may be applied to vaccinations to confer protection against infection with any bacteria having polysaccharide antigens, including, for example, Haemophilus influenzae, Streptococcus pneumoniae, Escherichia coli, Neisseria meningitidis, Salmonella typhi, Streptococcus mutans, Cryptococcus neoformans, Klebsiella, Staphylococcus aureus and Pseudomonas aeruginosa.

In another embodiment, the carrier function of D15
may be used, for example, to induce immunity toward
abnormal polysaccharides of tumor cells, or to produce

10

15

20

25

30

35

rD15 protein and its fragments were found to cross-react immunologically with the native D15 antigen isolated from both typeable and non-typeable H. influenzae isolates and thus represent cross-reactive immunogens for inclusion in a vaccine against diseases caused by H. influenzae. Furthermore, Haemophilus convalescent serum recognized D15 purified from H. influenzae as described herein, rD15 and N-terminal rD15 fragment.

another embodiment, the present invention provides a gene coding for the outer membrane protein D15 from H. influenzae having the specific nucleotide sequences described herein or ones substantially homologous thereto (i.e. those which hybridize under stringent conditions to such sequences), for genetically engineering hybrids or chimeric proteins containing a D15 fragment fused to another polypeptide or protein or a polysaccharide, such as H. influenzae outer membrane proteins, for example, P1, P2, or P6 or PRP. result, the hybrids, chimeric proteins or glycoconjugates may have higher protectivity against H. influenzae than D15, or P1, or P2, or P6, or PRP alone.

Thus, D15 outer membrane protein can function both as a protective antigen and as a carrier in a conjugate vaccine to provide autologous T-cell priming, wherein the of the conjugate hapten part is the capsular polysaccharide moiety (PRP) of H. influenzae. carbohydrate conjugate can elicit antibodies against both PRP and D15, and thus should enhance the level of H. influenzae-related diseases, protection against especially in infants.

In another embodiment, the present invention comprises an essentially pure form of at least one protein or peptide containing an amino acid sequence corresponding to at least one antigenic determinant of D15, which peptide is capable of eliciting polyclonal antibodies against <u>H. influenzae</u> in mammals. These D15-

10

20

25

<u>catarrhalis</u>, <u>Staphylococcus</u> <u>aureus</u>, or respiratory syncytial virus, in the presence or absence of adjuvant.

The D15 peptides (Table 2) or any portion, variant or mutant thereof, can easily be synthesized either manually or with a commercially available peptide synthesizer, such as the Applied Biosystems Model 430A synthesizer.

It is clearly apparent to one skilled in the art, that the various embodiments of the present invention have many applications in the fields of vaccination, diagnosis, and treatment of diseases caused <u>Haemophilus</u> infections, and the generation of immunological reagents. A further non-limiting discussion of such uses is further presented below.

15 1. Vaccine preparation and use

Immunogenic compositions, suitable for use as vaccines, may be prepared from immunogenic D15 outer membrane protein, fragments or analogs thereof and/or peptides corresponding to portions of D15 as disclosed The vaccine elicits an immune response which herein. produces antibodies, including anti-D15 outer membrane protein antibodies and antibodies against D15 that are opsonizing or bactericidal. Should the vaccinated subject be challenged by Haemophilus, the antibodies bind to the D15 outer membrane protein and thereby inactivate the bacterium. Opsonizing and bactericidal antibodies represent examples of antibodies useful in protection against disease.

Vaccines containing peptides are generally well
known in the art, as exemplified by U.S. Patents
4,601,903; 4,599,231; 4,599,230; and 4,596,792; all of
which references are incorporated herein by reference.
As to any further reference to patents and references in
this description, they are as well hereby incorporated by
reference without any further notice to that effect.
Vaccines may be prepared as injectables, as liquid

10

15

20

25

30

35

However, suitable dosage ranges are readily determinable by one skilled in the art and may be of the order of micrograms of the D15 outer membrane protein, analog, fragment and/or peptides. Suitable regimes for initial administration and booster doses are also variable, but may include an initial administration followed by subsequent administrations. The dosage of the vaccine may also depend on the route of administration and varies according to the size of the host.

The nucleic acid molecules encoding the D15 outer membrane protein of the present invention may also be used directly for immunization by administration of the DNA directly, for example, by injection for genetic immunization or by constructing a live vector, such as Salmonella, BCG, adenovirus, poxvirus or vaccinia. A discussion of some live vectors that have been used to carry heterologous antigens to the immune system are discussed in, for example, O'Hagan (1992). Processes for the direct injection of DNA into test subjects for genetic immunization are described in, for example, Ulman et al. (1993).

The use of peptides in vivo may first require their chemical modification since the peptides themselves may not have a sufficiently long serum and/or tissue halflife. Such chemically modified peptides are referred to herein as peptide analogs. The term peptide analog extends to any functional chemical equivalent of a peptide characterized by its increased stability and/or efficacy in vivo or in vitro in respect of the practice The term peptide analog is also used of the invention. herein to extend to any amino acid derivative of the described herein. peptides as Peptide analogs contemplated herein are produced by procedures that include, but are not limited to, modifications to side chains, incorporation of unnatural amino acids and/or their derivatives during peptide synthesis and the use of

10

nitration with tetranitromethane to form a 3-nitrotyrosine derivative.

Modification of the imidazole ring of a histidine residue may be accomplished by alkylation with iodoacetic acid derivatives or N-carbethoxylation with diethylpyrocarbonate.

Examples of incorporating unnatural amino acids and derivatives during peptide synthesis include, but are not limited to, use of norleucine, 4-amino butyric acid, 4-amino-3-hydroxy-5-phenylpentanoic acid, 6-aminohexanoic acid, t-butylglycine, norvaline, phenylglycine, ornithine, sarcosine, 4-amino-3-hydroxy-6-methylheptanoic acid, 2-thienyl alanine and/or D-isomers of amino acids.

2. Immunoassays

The D15 outer membrane protein, analog, fragment 15 and/or peptides of the present invention are useful as antigens in immunoassays, including enzyme-linked immunosorbent assays (ELISA), RIAs and other non-enzyme linked antibody binding assays or procedures known to the 20 art for the detection of anti-bacterial, Haemophilus, D15 and/or peptide antibodies. In ELISA assays, the D15 outer membrane protein, fragment or analogs thereof and/or peptides corresponding to portions of D15 outer membrane protein are immobilized onto a selected surface, 25 for example, a surface exhibiting a protein affinity, such as the wells of a polystyrene microtiter plate. After washing to remove incompletely adsorbed D15 outer membrane protein, analog, fragment and/or peptides, a nonspecific protein, such as bovine serum albumin (BSA) 30 or casein, that is known to be antigenically neutral with regard to the test sample may be bound to the selected This allows for blocking of nonspecific adsorption sites on the immobilizing surface and thus decreases the background caused by nonspecific bindings 35 of antisera onto the surface. Normally, the peptides

10

15

20

25

30

35

<u>Haemophilus</u> and other bacteria that have genes encoding D15 outer membrane proteins.

The nucleotide sequences comprising the sequence encoding the D15 outer membrane protein of the present invention are useful for their ability to selectively form duplex molecules with complementary stretches of other D15 genes. Depending on the application, a variety of hybridization conditions may be employed to achieve varying degrees of selectivity of the probe toward the other D15 genes. For a high degree of selectivity, stringent conditions are used to form the duplexes, such as low salt and/or high temperature conditions, such as provided by 0.02 M to 0.15 M NaCl at temperatures of between about 50°C to 70°C. For some applications, less stringent hybridization conditions are required such as 0.15 M to 0.9 M salt, at temperatures ranging from between about 20°C to 55°C. Hybridization conditions can also be rendered more stringent by the addition of increasing amounts of formamide, to destabilize the hybrid duplex. Thus, particular hybridization conditions can be readily manipulated, and will generally be a method of choice depending on the desired results.

In a clinical diagnostic embodiment, the nucleic acid sequences of the D15 outer membrane protein genes of the present invention may be used in combination with an appropriate means, such as a label, for determining hybridization. A wide variety of appropriate indicator means are known in the art, including radioactive, enzymatic or other ligands, such as avidin/biotin, which are capable of providing a detectable signal. In some diagnostic embodiments, an enzyme tag, such as urease, alkaline phosphatase or peroxidase, instead of a radioactive tag may be used. In the case of enzyme tags, colorimetric indicator substrates are known which can be employed to provide means visible to the human eye or

10

15

20

25

30

35

The pBR322 plasmid, or other microbial plasmid or phage must also contain, or be modified to contain, promoters which can be used by the microbial organism for expression of its own proteins.

In addition, phage vectors containing replicon and control sequences that are compatible with the host microorganism can be used as a transforming vector in connection with these hosts. For example, the phage in lambda $GEM^{TM}-11$ may be utilized in making recombinant phage vectors which can be used to transform host cells, such as E. coli LE392.

Promoters commonly used in recombinant DNA construction include the β -lactamase (penicillinase) and lactose promoter systems and other microbial promoters, such as the T7 promoter system. Details concerning the nucleotide sequences of promoters are known, enabling a skilled worker to ligate them functionally with plasmid The particular promoter used generally is a matter of choice depending upon the desired results. Hosts that are appropriate for expression of transferrin receptor genes, fragment analogs or variants include E. coli, Bacillus, Haemophilus, Bordetella, fungi, yeast, or the baculovirus and poxvirus expression systems may be used.

In accordance with an aspect of this invention, it is preferred to make the D15 outer membrane protein, fragment or analog thereof by recombinant methods, particularly since the naturally occurring D15 protein as purified from culture of a species of Haemophilus may include undesired contaminants, including trace amounts of toxic materials. This problem can be avoided by using recombinantly produced D15 outer membrane protein in heterologous systems which can be isolated from the host in a manner to minimize toxins in the purified material. Particularly desirable hosts for expression in this regard include Gram positive bacteria which do not have

10

20

25

30

35

the invention. Although specific terms have been employed herein, such terms are intended in a descriptive sense and not for purposes of limitations. Immunological and recombinant DNA methods may not be explicitly described in this disclosure but are well within the scope of those skilled in the art.

EXAMPLES

Methods of molecular genetics, protein biochemistry, and immunology used but not explicitly described in this disclosure and these EXAMPLES are amply reported in the scientific literature and are well within the ability of those skilled in the art.

Example 1

This Example illustrates the cloning and sequencing of the D15 genes.

Genomic DNA was purified from the <u>Haemophilus</u> influenzae type b strain Ca by lysis of the bacteria with pronase and sodium dodecylsulphate followed by phenol extraction and isopropanol precipitation, according to Berns and Thomas, 1965. The DNA was then partially digested with EcoRI and the DNA fraction containing 6-10 kb fragments was isolated following electrophoresis in low-melting point agarose. These fragments were ligated into a lambda gtll Ampl vector (Thomas and Rossi, 1986) and cloned as a lysogen into E. coli strain BTA282. Recombinant clones were selected for their ampicillin resistance conferred by the vector. To identify clones producing H. influenzae type b antigen, the clones were replica-plated on nitrocellulose filters and duplicate colonies induced for expression by temperature switch to 42°C for 2 hours. Colonies were lysed by wetting the filters with 1% sodium dodecylsulphate (SDS). The filters were then placed into a chloroform-saturated atmosphere for 15 min. The filters were then assayed by colony radioimmuno-assay using a hyperimmune rabbit anti-H. influenzae type b antiserum absorbed with E. coli lysate

15

20

25

30

plasmid transformed into <u>E. coli</u> HB101. Recombinant bacteria were found to produce the expected M_r 80 kDa <u>H. influenzae</u> type b antigen when examined by Western blotting. The insert DNA was then characterised by restriction endonuclease mapping. A 2.8 kb <u>HindIII-Eco</u>RI fragment was subcloned into pUC19 to generate plasmid pUC19/D15, which was transformed into <u>E. coli</u> HB101. The recombinant bacteria expressed a M_r 80 kD protein recognized by D15-specific antibodies on Western blot analysis of <u>E. coli</u> lysates.

Plasmid DNA was prepared from two individual colonies of recombinant E. coli HB101 containing the pUC19/D15 plasmid using standard 'techniques. Oligonucleotide sequencing primers of 17-25 bases in length were synthesized on the ABI model Synthesizer and purified by chromatography using OPC cartridges obtained from Applied Biosystems Inc., used in accordance with the manufactures recommendations. Samples were sequenced using the ABI model 370A DNA Sequencer and dye terminator chemistry according to manufacturers' protocols. This sequence indicated that the D15 gene contains an open reading frame encoding for 789 amino acids, including a putative signal sequence (Figure 1). The derived amino acid sequence was found to contain the sequence of an internal peptide obtained by thrombin digestion of native D15 that had been chemically determined. The amino composition of D15 derived from the D15 gene sequence was comparable (within experimental error) to that of the native protein as determined by amino acid analysis.

Example 2

This Example illustrates the preparation of chromosomal DNA from <u>Haemophilus influenzae</u> strains Eagan, MinnA, SB33, and PAK 12085.

35

H. influenzae Eagan and PAK 12085 chromosomal DNAs were digested with Sau3A I (0.5 unit/10 μ g DNA) at 37°C for 15 minutes and size-fractionated by agarose gel electrophoresis. Gel slices corresponding DNA fragments of 15-23 kb were excised and DNA electroeluted overnight in dialysis tubing containing 3 mL of TAE (40mM Tris-acetate, 1mM EDTA, pH 8.0) at 14V. The DNA was precipitated twice and resuspended in water before overnight ligation with EMBL3 BamH I (Promega). The ligation mixture was packaged using the Lambda in vitro packaging kit (Amersham) according to the manufacturer's instructions and plated onto E. coli NM539 The library was titrated, then amplified and stored at 4°C under 0.3% chloroform.

MinnA chromosomal DNA (10 μ g) was digested with 15 Sau3A I (40 units) for 2, 4, and 6 minutes then sizefractionated on a 10-30% sucrose gradient in TNE (20mM Tris-HCl, 5mM NaCl, 1mM EDTA, pH 8.0). Fractions containing DNA fragments >5 kb were pooled 20 precipitated. In a second experiment, chromosomal DNA (2.6 μ g) was digested with Sau3A I (4 units) for 1, and 3 minutes and size- fractionated by preparative agarose gel electrophoresis. Gel slices containing DNA fragments of 10-20 kb were excised and DNA extracted by 25 a standard freeze/thaw technique. The size-fractionated DNA from the two experiments was pooled for ligation with BamH I arms of EMBL3 (Promega). The ligation mix was packaged using the Gigapack II packaging kit (Amersham) and plated on E. coli LE392 cells. The library was titrated, then amplified and stored at 4°C under 0.3% 30 chloroform.

SB33 chromosomal DNA (20 μ g) was digested with Sau3A I (40 units) for 2, 4, or 6 minutes and size-fractionated on a 10-30% sucrose gradient in TNE (20mM Tris-HCl, 5mM NaCl, 1mM EDTA, pH 8.0). Fractions containing fragments >5 kb were pooled. In a second experiment, SB33

10

15

20

25

compared with the amino acid sequence of the D15 protein of <u>H. influenzae</u> type b Ca (Figure 3). Example 5

This Example illustrates the expression of rD15 protein in E. coli.

A 2.8 kb fragment <u>Hind</u>III-<u>Eco</u>RI was subcloned into pUC19 and this pUC19/D15 plasmid was transformed into E. coli HB101. Upon induction, the positive clones expressed an 80 kDa protein which was recognized by D15-specific antisera on Western blot analysis. A <u>Hind</u>III-Pst I fragment was also subcloned into pUC19 and shown to express a 67 kDa protein. According to the restriction map, this 67 kDa protein corresponded to a C-terminal truncated D15 protein. On Western blot analysis, this truncated D15 was still recognized by the D15-specific antisera.

Plasmids to express the D15 gene of the non-typeable strain SB33 in E. coli were constructed. Plasmid JB-1042-5-1 containing the SB33 D15 gene and its flanking regions, was digested with <a>EcoR I and <a>Hind III and the 3kb D15 insert subcloned into pUC to give plasmid pRY-60-1 (Figure 4). Appropriate oligonucleotides were synthesized to restore the native D15 sequence between the ATG codon of the expression plasmid pT7-7 and the BsrF I site within the D15 gene. These oligonucleotides had the following sequence:

Nde

5'- TATGGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGGTGTTCAAGGTG 30 ACCGTGGAAAACACCGTTTTCTATAAGCACACCTACCACAAGTTCCACTGAATCT ACTTAGAATCAACAAACCGAGCAAGTTTACCTGTTCGTG - SEQ ID NO: 50 35 TGGTTGTTTAGGCTCGTTCAAATGGACAAGCACGGCC-5'- SEQ ID NO: 51

BsrF I

Plasmid pRY-60-1 was digested with EcoR I and BsrF I and the DNA fragment containing most of the D15 gene was 40 purified. pUC was digested with EcoR I and Nde I and the

10

15

25

30

35

produced by transformed \underline{E} , $\underline{\operatorname{coli}}$ was isolated by affinity purification on glutathione agarose.

Example 7

This Example describes alternative expression systems for rD15.

The D15 gene or fragments thereof are also expressed in E. coli under the control of other regulated promoters. The D15 gene or fragments thereof are expressed in the absence of the leader peptide, or in other cloning systems where toxicity of D15 expression to the host is not problematic. The gene or fragments thereof are synthesized de novo or by employing the polymerase chain reaction using suitable primers. genes are cloned into suitable cloning vectors or bacteriophage vectors in E. coli or other suitable hosts directly when toxicity can be avoided. Expression systems are Gram-positive bacteria (such as Bacillus species), pox virus, adenovirus, baculovirus, yeast, fungi, BCG or mammalian expression systems.

20 Example 8

This Example illustrates the protocol for extraction and purification of rD15 from $\underline{E.\ coli}$ expression system.

The cell pellet from a 250 mL culture, prepared as described in Example 5, was resuspended in 40 mL of 50 mM Tris, pH 8.0, and disrupted by sonication (3 x 10 min, 70% duty circle). The extract was centrifuged at 20,000 x g and the resulting pellet saved. The initial pellet was re-extracted with 40 mL of 50 mM Tris, 0.5% Triton X-100, 10 mM EDTA, pH 8.0. The suspension was then sonicated for 10 minutes at 70% duty circle. The extract was centrifuged at 300 x g for 5 minutes. The resulting supernatant was centrifuged again at 20,000 x g for 30 min and the resulting pellet was saved. The pellet was resuspended in 50 mM Tris, 0.5% Triton X-100, 10 mM EDTA, pH 8.0. The suspension was then mixed with PBS/ 8 M urea to a final urea concentration of 6 M. The solution was

10

15

20

25

30

35

This Example illustrates the procedure used for N-terminal rD15 fragment purification from GST using Glutathione-Sepharose 4B affinity chromatography.

A thrombin-digested GST-(D15 fragment) sample, prepared as described in Example 10, was loaded onto a Glutathione-Sepharose 4B column (2 mL) equilibrated with PBS containing 1% Triton X-100. The run-through of the column containing the N-terminal rD15 fragment was saved. After washing the column with 20 mL of PBS, the affinity column was regenerated by removing GST using 50 mM Tris-HCl buffer, pH 8.0, containing 5 mM glutathione. The purity of rD15 fragment was analysed by SDS-PAGE (Figure 9, lane 5). This N-terminal rD15 fragment contains amino acids 63-223 of the D15 protein as a result of cleavage at the spacious thrombin site shown in Figure 1A.

Example 12

This Example illustrates the protocol used for the purification of D15-specific polyclonal antibodies by affinity chromatography using GST-(D15 fragment) fusion protein.

The recombinant GST-(D15 fragment) fusion protein, prepared as described in Example 9, was conjugated to cyanogen bromide-activated Sepharose. The affinity column was then used to purify antibodies from a rabbit hyperimmune anti-H. influenzae type b antiserum. The affinity purified-antibodies were shown by immunoblotting to react with a 80 kDa component present in the lysates of E. coli transformed with pUC9/D15 and in the lysates of several typeable and nontypeable H. influenzae isolates. These results confirmed that the DNA segment encoding the D15 fragment of the fusion protein was part of the open reading frame of the D15 gene.

Similarly, antisera raised against the recombinant fusion protein (Example 9) or the purified N-terminal rD15 fragment (Example 11) reacted with the D15 protein produced by H. influenzae strains (Example 13).

15

20

25

30

35

in Example 17. The mean molecular size of the PRP molecules used for conjugation was determined as being approximately 20,000 Daltons. The conjugation was carried out without a linker molecule but may also be carried out with a linker molecule. A PRP/D15 molar ratio of approximately 7 was used to provide an excess of PRP hapten.

The PRP/rD15 conjugate was tested according to the protocol of Example 18 for immunogenicity in rabbits and elicited both primary and secondary anti-PRP IgG and anti-D15 antibody responses (Table 9). Rabbit anti-rD15-PRP antisera also strongly reacted with both native D15 and rD15 as judged by immunoblot analysis. These data indicate that rD15 can be used as a carrier protein in a conjugate vaccine. In addition, a rD15-PRP conjugate vaccine should ensure a more consistent protection against H. influenzae type b disease, particularly in infants, as a result of the additional homotypic protection provided by antibodies directed against the D15 protein.

Example 15

This Example describes the preparation of D15 peptides.

D15 peptides (Table 2) were synthesized using an ABI 430A peptide synthesizer and optimized t-Boc chemistry as described by the manufacturer, then cleaved from the resin by hydrofluoric acid (HF). The peptides were purified by reversed-phase high performance liquid chromatography (RP-HPLC) on a Vydac C4 semi-preparative column (1 x 30 cm) using a 15 to 55% acetonitrile gradient in 0.1% trifluoryl acetic acid (TFA) developed over 40 minutes at a flow rate of 2 mL/min. All synthetic peptides (Table 2) used in biochemical and immunological studies were >95% pure as judged by analytical HPLC. Amino acid composition analyses of

10

15

25

30

35

equilibrated with 0.2 M sodium phosphate buffer, pH 7.2, and eluted with the same buffer. Fractions were monitored for absorbance at 230 nm. The first major protein peak was pooled and concentrated in a Centriprep 30 to 2.2 mL. The amount of protein was determined using the Bio Rad protein assay, and was found to be 300 μ g/mL. The presence of PRP in the protein conjugate fraction was confirmed by the Orcinol test.

Example 18

This Example describes the protocol used for the production of anti-PRP antisera in animals using rD15-PRP conjugates.

Rabbits were immunized intramuscularly with rD15-PRP conjugates (Example 14) (5 to 50 μ g PRP equivalent) mixed with 3 mg AlPO₄ per mL, followed by two booster doses (half amount of the same immunogen) at 2 week intervals. Antisera were collected every 2 weeks after the first injection, heat-inactivated at 56°C for 30 minutes and stored at -20°C.

20 Example 19

This Example illustrates the reactivity between D15 peptides and anti-peptide and D15-specific antisera using D15-specific and peptide-specific ELISAs.

Microtiter wells (Nunc-Immunoplate, Nunc, Denmark) were coated with 200 ng of purified rD15 or 500 ng of individual peptides in 50 μ L of coating buffer (15 mM Na₂CO₃, 35 mM NaHCO₃, pH 9.6) for 16 hours at room temperature. The plates were then blocked with 0.1% (w/v) BSA in phosphate buffer saline (PBS) for 30 minutes at room temperature. Serially diluted antisera were added to the wells and incubated for 1 hour at room temperature. After removal of the antisera, the plates were washed five times with PBS containing 0.1% (w/v) Tween-20 and 0.1% (w/v) BSA. F(ab')₂ fragments from goat anti-rabbit, guinea pig, mouse, or human IgG antibodies conjugated to horseradish peroxidase (Jackson ImmunoResearch Labs Inc.,

15

20

25

30

35

the substrate tetramethylbenzidine (TMB) in $\rm H_2O_2$ (ADI, Toronto). The reaction was stopped with 1N $\rm H_2SO_4$ and the optical density measured at 450 nm using a Titretek Multiskan II (Flow Labs., Virginia). A standard anti-PRP antiserum of known titer was included as positive control. Assays were performed in triplicate, and the reactive titer of each antiserum was defined as the reciprocal of the dilution consistently showing a 2-fold increase in O.D. value over that obtained with the pre-immune serum (Table 9).

Example 21

This Example describes the protocol used for the production of D15-specific antisera using purified D15, rD15 or N-terminal rD15 fragment.

New Zealand White rabbits (Maple Lane) and guinea pigs (Charles River) were immunized intramuscularly (IM) with a 10 µg dose of either affinity-purified native D15 (Example 13), recombinant D15 (Example 8) or N-terminal rD15 fragment (Example 11) emulsified in Freund's complete adjuvant (Difco). Animals were boosted on day 28 with another 10 µg dose of affinity-purified D15 or rD15 or rD15 fragment emulsified in Freund's incomplete adjuvant and bled on day 42 via the marginal ear vein. D15-specific polyclonal antibodies were purified from this material as described in Example 12.

Example 22

This Example illustrates the protective activity of D15-specific antisera against <u>H. influenzae</u> type b challenge using the infant rat model of bacteremia.

Five-day old infant rats were inoculated subcutaneously (SC) on the dorsum with 0.15 mL of two different rabbit anti-N-terminal rD15 fragments. Pre-immune sera were used as negative controls. One day after immunization, the infant rats were injected intraperitoneally (IP) with 200 colony-forming units (cfu) of Haemophilus influenzae type b Minn A strain (0.1)

10

15

20

25

30

35

supernatant added to further expand and maintain the viability of the peptide-specific T-cells. After a further 6 day-incubation, the cells were washed three times, each time with 200 μL of culture medium.

Each set of cultures was then stimulated with the corresponding concentrations (1, 10 and 100 µg per mL) of the peptide in the presence of 2×10^5 irradiated (1,500) rad) BALB/c spleen cells in a final volume of 200 µL of culture medium. Sixty µL of supernatant were then removed from each microculture. The supernatants from each triplicate cultures set were pooled. All supernatants were assayed for IL-2, Interleukin-4 and Interferon-gamma Detections of IL-2 and IL-4 were performed using murine IL-2 and IL-4 ELISA kits purchased from Endogen Inc. (MA, USA) respectively. Assay of IFN- γ was performed using a mouse IFN-γ ELISA kit supplied by Genzyme Corporation (MA, USA). Test culture supernatants were assayed at 1 in 5 dilution according to the manufacturers' instructions. The results obtained are set forth in Table 7.

Example 25

This Example describes the general procedure used for the production of murine D15-specific monoclonal antibodies.

BALB/c mice were immunized intraperitoneally with 20 to 50 μg of the N-terminal rD15 fragment (Example 11) emulsified in Freund's complete adjuvant. Two weeks later, the mice were given another injection of the same amount of immunogen in incomplete Freund's adjuvant (IFA). Three days before the fusion, the mice were boosted again with the same amount of immunogen in IFA. Hybridomas were produced by fusion of splenic lymphocytes from immunized mice with non-secreting Sp2/0 myeloma cells as previously described by Hamel et al. (1987). D15-specific hybridomas were cloned by sequential limiting dilutions and screened for anti-D15 monoclonal

TABLE 1

PROTECTIVE EFFECT OF PASSIVELY TRANSFERRED ANTI-N-TERMINAL RD15 FRAGMENT ANTIBODIES IN THE INFANT RAT MODEL OF BACTEREMIA'

cfu/0.1 mL blood					
Rabbit antisera	Pre-immune	Post-immunization	p value		
Rb#434	510 (6/6)2	6 (1/6)	<0.001		
Rb#435	910 (4/4)	6 (1/4)	<0.001		

Five-day old infant rats were passively immunized with 0.15 mL of rabbit anti-N-terminal rD15 fragment s.c. One day later, the infant rats were challenged with 200 cfu of $\underline{\text{H. influenzae}}$ type b strain MinnA (0.1 mL, IP). The blood samples were taken from each rat 24 hours after the challenge and analysed for bacteria counts.

The parentheses indicate the number of rats found to be bacteremic out of the total number of rats challenged.

D15-P29	619-646	LWVVSAKASAGYANGFGNKRLPFYQTYT	42
D15-P30	641-666	fyqtytaggigslrgfaygsigpnai	43
D15-P31	662-688	GPNAIYAEYGNGSGTGTFKKISSDVIG	44
D15-P32	681-709	KISSDVIGGNAIATASAELIVPTPFVSDK	4.5
D15-P33	705-731	FVSDKSQNTVRTSLFVDAASVWNTKWK	46
D15-P34	725-750	VWNTKWKSDKNGLESDVLKRLPDYGK	47
D15-P35	745-771	LPDYGKSSRIRASTGVGFQWQSPIGPL	4.8
D15-P36	769-798	GPLVFSYAKPIKKYENDDVEOFOFSIGGSF	4 9

TABLE 4

INHIBITION OF ANTI-N-TERMINAL rD15 FRAGMENT ANTIBODY-INDUCED PROTECTION BY D15 PEPTIDES IN THE INFANT RAT MODEL OF BACTEREMIA

Group #	Antibody	cfu / 10 μL blood	cfu in each group/ cfu in group #4 (control) (%)
1	Anti-D15 Ab + PBS	$60 \pm 120 (3/7)$	3
2	Anti-D15 Ab + peptides	300 ± 240 (6/7)	. 13
3	Anti-D15 Ab + rD15	1,520 ± 1,280 (7/7)	64
4	PBS + peptides	2,360 ± 1,200 (6/7)	100

One half mL of rabbit anti-N-terminal rD15 fragment antiserum (Anti-rD15 fragment Ab) was mixed with either nine D15 peptides (100 $\mu \rm g$ of peptides D15-P2 to D15-P10, See TABLE 2) or with 600 $\mu \rm g$ of N-terminal rD15 fragment at room temperature for 1 hr. Antiserum and peptides mixed with PBS were used as controls. Seven-day old infant rats were injected s.c. with 0.2 mL of the various preparations. After 24 h, the infant rats were challenged I.P. with 200 cfu of H. influenzae type b strain MinnA. The blood samples were taken at 24 h after the challenge. The numbers in parentheses indicate the number of animals that were bacteremic out of the total number of animals challenged. The level of bacteremia is expressed as the mean of values obtained from seven infant rats tested individually \pm one standard deviation (SD).

TABLE 6

REACTIVITY OF RABBIT, GUINEA PIG AND MOUSE ANTI-rD15 ANTISERA WITH D15 PEPTIDES

Peptide	Rabbit ²	Reactive Titer	
repute	KADDIT -	Guinea Pig ³	Mouse
D15-P1	-	•	+
D15-P2	· •	+++	+,
D15-P3	-	-	+
D15-P4	+	+	+
D15-P5	•	• · .	+
D15-P6	-	+	+
D15-P7	-	•	+
D15-P8	-	++++	++++
D15-P9	•	-	+
D15-P10		-	+++
D15-P11	-	=	+++
D15-P12	-	-	+
D15-P13	-	•	+
D15-P14	+++	+	+
015-P15	-	-	+
015-P16	-	• •	+
015-P17	-	-	+
015-P18	-	•	+
)15-P19	-	•	+
015-P20	-	•	+
15-P21	-	•	+
15-P22	-	•	+
15-P23	-	, •	+
15-P24	-	-	+
15-P25	•	-	+
15-P26	-	•	+++
15-P27	• .	. +	+

TABLE 7
T-CELL STIMULATORY ACTIVITY OF D15 PEPTIDES

D15-P6	Peptide	IL-2 ²	CYTOKINE RELEASE (pg/mL) γ-IFN³	IL-4 ⁴
D15-P3	D15-P1	-		-
D15-P4 -	D15-P2	122	-	-
D15-P5 742 38,000 1 D15-P6	D15-P3	25	•	-
D15-P6	D15-P4	- .		_
D15-P7	D15-P5	742	38,000	13
D15-P8	D15-P6	-	-	-
D15-P9 -	D15-P7	-	-	-
D15-P10 108 1,900 - D15-P11	D15-P8	-	*	- -
D15-P11	D15-P9	-	• •	-
D15-P12	D15-P10	108	1,900	-
D15-P13	D15-P11	-	-	-
D15-P14	D15-P12	1,052	6,100	٠_
D15-P16	D15-P13	105	6,200	56
D15-P16 48	D15-P14	-	-	-
D15-P17	D15-P15	-	-	-
D15-P18 32 4,800 _ D15-P19 882 24,500 _ D15-P20 - _ _ D15-P21 - _ _ D15-P22 - _ _ D15-P23 78 _ _ D15-P24 103 _ _ D15-P25 - _ _ D15-P26 572 6,700 _	D15-P16	48	· <u>-</u>	-
D15-P19 882 24,500	D15-P17	•	-	•
D15-P20	D15-P18	32	4,800	- ,
D15-P21	D15-P19	882	24,500	-
D15-P22	D15-P20	-	-	. -
D15-P23 78 D15-P24 103 D15-P25 D15-P26 572 6,700 .	D15-P21	-	-	-
D15-P24 103 D15-P25	D15-P22	-	- '	•
D15-P25 D15-P26 572 6,700 -	D15-P23	78	-	•
D15-P26 572 6,700 -	D15-P24	103	-	-
	D15-P25	-	- .	-
D15-P27 274 7,505 68	D15-P26	572	6,700	-
	D15-P27	274	7,505	68

67

TABLE 8

RABBIT AND GUINEA PIG ANTIBODY RESPONSES TO D15 PEPTIDES

	Peptide-s	Peptide-specific ELISAs		
_	Reactive	Titer ¹		
Immunogen	Rabbit ²	Guinea Pig³		
D15-P1	102,400	819,200		
D15-P2	204,800	1,637,400		
D15-P3	51,200	1,637,400		
D15-P4	204,800	819,200		
D15-P5	51,200	1,637,400		
D15-P6	51,200	409,600		
D15-P7	204,800	819,200		
D15-P8	51,200	409,600		
D15-P9	102,400	409,600		
015-P10	102,400	819,200		
015-P11	51,200	819,200		
015- P1 2	102,400	204,800		
015-P13	NT4	204,800		
015-P14	NT	409,600		
15-P15	NT	204,800		
P15-P16	NT	819,200		
015-P17	NT	204,800		
015-P18	NT	312,500		
015-P19	NT	312,500		
015-P20	NT	62,500		
15-P21	NT	62,500		
015-P22	NT	12,500		
15-P23	NT	1,562,500		
015-P24	NT	312,500		
15-P25	NT	62,500		

TABLE 9

RABBIT IGG ANTIBODY RESPONSE TO D15-PRP CONJUGATE

		Reactive	Titer Against	;²
Rabbit¹ #	PRP		rD15	
	2 doses	3 doses	2 doses	3 doses
489-1	1,600	3,200	1,600	6,400
490-1	1,600	1,600	6,400	25,600

Rabbits were immunized intramuscularly with rD15-PRP conjugates (5 to 50 μg PRP equivalent) mixed with 3 mg ALPO, per mL, followed by two booster doses (half amount of the same immunogen) at 2 week intervals.

Reactive titres is based on PRP specific and D-15 specific ELISAS.

REFERENCES

O'Hagan (1992).

Ulman et al., 1993.

Berns C.A. and Thomas C.A. (1965) J. Mol. Biol. 11:476-490.

Thomas W.R. and Rossi A.A. (1986) Infect. Immun. **52**:812-817.

Thomas W.R. et al. (1990) Infect. and Immun. 58:1090-1913.

Carlone G.M. et al. (1986) J. Clin. Microbiol. 24:330-331.

Smith, D.B. and Johnson K.S. (1988) Gene 67:31-40.

Harkness, R. et al. (1992) J. Bacteriol. 174:2425-2430.

Hamel et al. (1987) J. Med. Microbiol. 23:163-170.

Mills et al. (1993) Infect. Immun. 61:399-410.

Trinchieri, (1993) Immunol. Today 14:335-338.

Hope, T.P. (1986) J. Immunol. Methods 88:1-18.

Zangwill et al., 1993. MMWR 42:1-15.

Loeb et al. 1987. Infect. Immun. 55:2612-2618.

Panezutti, 1993. Infect. Immun. 61: 1867-1872.

- 8. The recombinant vector of claim 6 wherein said DNA segment encodes a polypeptide of at least 6 residues.
- 9. The recombinant vector of claim 8 wherein said polypeptide is selected from those shown in Table 2.
- 10. The recombinant vector of claim 6, 7, 8 or 9 wherein said DNA segment consists of no more than the coding sequence for said D15 outer membrane protein.
- 11. The recombinant vector of claim 10, wherein the DNA segment further comprises a nucleic acid sequence encoding a leader sequence for export of said gene product from said host.
- 12. A purified and isolated protein encoded by the DNA fragment contained in the recombinant vector of claim 10 or 11.
- 13. A purified and isolated D15 outer membrane protein, or a portion thereof.
- 14. The protein of claim 13 wherein the D15 outer membrane protein is a <u>Haemophilus</u> D15 outer membrane protein.
- 15. The protein of claim 14 wherein the D15 outer membrane protein is a <u>Haemophilus influenzae</u> D15 outer membrane protein.
- 16. The protein of claim 15 wherein the <u>Haemophilus</u> influenzae is a type b <u>Haemophilus influenzae</u> strain.
- 17. The protein of claim 16 wherein the <u>Haemophilus</u> influenzae type b strain is selected from Ca, MinnA and Eagan strains.
- 18. The protein of claim 15 wherein the <u>Haemophilus</u> influenzae is a non-typeable <u>Haemophilus</u> influenzae strain.
- 19. The protein of claim 18 wherein the non-typeable Haemophilus influenzae strain is selected from PAK12085 and SB33 strains.
- 20. A synthetic peptide containing an amino acid sequence corresponding to the amino acid sequence of the protein or portion thereof claimed in any one of claims

- 30. The chimeric molecule of claim 29 wherein said another polypeptide or protein comprises a P1, P2 or P6 outer membrane protein of <u>H. influenzae</u>.
- 31. The chimeric molecule of claim 28 wherein said polysaccharide comprises a PRP molecule from $\underline{\text{H.}}$ influenzae.

FIG.1A

H. influenzae b Ca strain D15 sequence

		1/82			•
T 09	G A 120	C T 180	I A T.G 240	G C 300	T A 360
T T A	THR A·C G	ASP LGACT	ASN A A T (241	ALA G C G	WAL G T T
AT	THR A C A A	GLY G G T (ASP 3 A C	LYS A A A	ASP VAL L SATGTTA 360
-10 A T A	THR A		THR A C T G	VAL 3 T G A	SER CAC
. T. 02	GLY 1 3 G T A 110	GLY VAL GLA 3 G T G T T C A A 170	VAL 1 3 T G A 230	ASP 7 3 A T G 290	ILE :
D C	មិ ខេត	VA I G II	VA 1 G T	AS I G A	
T A (雅 T T C (ARG C G T (ASP G A T	ILE A T C A
l l	LEU TTA	ASP GLY VAL GATGGTGTT 170	GLY GIN ARG VAL THR ASP ASN 13 GTCAGCGTGTGACTGACAATG	PE TTC	LYS SER AAATCGA 140
АТТ 40	LEU TTA	ARG VAL ASP GTGTGGAT 160	GLY G G T 220	ARG C G A 1 280	LYS A A A 340
-35 G <u>тт G C A</u> ттаттаат G A ттттта C G T C <u>T A T A A T</u> T T A T 30 40 50	SER AGT	()	ALA 3 C C	GLY S.G. T	IS GIN GLU GLY ASP VAL LEU VAL VAL SER VAL VAL ALA LYS SER ILE ILE SER ASP VAL 1 A T C A A G A A G G C G A T G T T G T T A G C G T T G T T G A T G T T T C A G A T G T T A 310 320 330 330
TTA	ALA G C A	ASP ILE SATATT(ARG CGT(SER AGTO	SER VAL VAL GCGTTGTG 330
гтат 30	ILE AATC 90	ASP G A T 50) VAL TGTT 210	C G T A 270	C G T T 330
CAT	LEU CTA	LYS A A A A G 150	PRO C C T Q 210	PHE TTC 27	SER AGC
-35 T T G	LEU CTT	ALA G C A	LEU TTA	LEU	VAL G T T
	LYS A A A	AL T G	SER AGTT	SER SER TO T	VAL G T T (
C G G 20	LYS A A A 80	T T T 140	ALA G C A 200	ARG C G C 260	LEU CTT 320
T A A	MET A T G	ALA ALA PRO PHE S C G C A C C T T T T 130 140	ARG ALA SER LES CGAGCAAGTTT 200 courious thrombin site	ILE VAL ARG TTGTCCGCT 260	ASP VAL BATGTGO
Hind III AGCT	TCG	ALA G C A	ILE A T C	ILE ATT	ASP GAT
Hind III GATTACGCCAAGCTTAACGGT 10	RBS ATAGGATACAATC 70	HR VAL PHE ALA ALA PRO PHE V CTGTGTTTGCCGCACCTTTG 130	EU GLU GIN GIN ILE ARG ALA TAGAACAACAAATCCGAGC1 190 comigus the	ASN A A T 1 250	IS GIN GLU GLY ATCAAGAAGGCO
ນ ၅ ນ	A 'F A	. Limicaled VAL PHE	GIN C A A	VAL ALA 3 T G G C T 7	GLU GAA
r T A	RBS A G G	VAL GTG	G A A	VAL G T G	GIN C A A
G A C	AT	H H C I G	EU T A	SP A T G	IS A T

			2/82			
ဗ် ဗ ဥ	A A 480	C C C 540		^ອ ອ	G C G 720	I 4
PRO THR GLU ALA LEU LYS GLN ASN LEU ASP ALA ASN G CCCACTGAAGCACTTAAACAAACTTAGATGCTAACG 390 420	VAL L GTAA 480	ARG TYR ASN ALA THR VAL GLU PRO ILE VAL ASN THR L CGCTATAACGCAACAGTTGAACCTATTGTCAATACGC 510 520 530	LEU A Trgg	ASN GLU SER VAL SER SER THR LEU GLN GLU GLN MET G A A C G A A T C T G T T A C A T T A C A A G A A A T G G 630 640 660) T. T. (ALA LYS ALA GIN I CCAAAGCACAAA 770
ALA G C T	SER A G T	ASN A A T	ALA LYS	GLN C A A	GIN C A A '	ALA G C A
ASP GATG	LYS A A A A	VAL GTC	ALA G C A	GLU GAA	ALA G C G	LYS A A A
LEU TTA 410	ALA G C C A 470	1LE A T T 530	LYS A A A G 590	GLN C A A 650	GLY G G T G 710	ALA G C C 770
ASN A A C	PHE TTT	PRO C C T	ASP ASP BATGATA	leu T T A	GLU GAA	TYR T A T
LYS GLN ASN A A C A A A A C C	ASN GLU PHE AATGAATTT 160	GLU GAA	ASP G A T	SER THR LEU GLN GLU GLN NGTACATTACAAGAACAA 840 650	PHE TTT	ASN GLY TYR AATGGCTATC 760
LYS A A A 400	ASN A A T 460	TYR ASN ALA THR VAL GLU PRO ILE TATAACGCAACAGTTGAACCTATTG 510 520	GLU G A A 580	SER A G T 640	LEU TRP GLY ASN LYS PHE GLU FTATGGGGAAATAAATTGAA(690	ASN A A T 760
PRO THR GLU ALA LEU CCACTGAAGCACTTA 390	LEU TTA1	THR ACA	ASN A A T	SER AGCA 6	ASN A A T	ASN A A T A
ALA G C A	LYS A A A '	ALA G C A	ILE A T C	SER AGTP	GLY G G A	ASP TYR TYR LEU A T T A T T A T T T A 1 750
GLU GAA 90	ARG CLU CAGAA 450	ASN TAAC 510	E GLN T C A A 570	SER VAL CTGTT 630	TRP T G G 30	TYR T A T 50
A C T	ARG C G A	TYR TAT 5	LEU ILE GIN TAATTCA!	SER TCT 6.	LEU TTA	TYR T
PRO C C C	ILE A T T	ARG C G C	LEU TTA	CLU GAA	LYS A A A	ASP GAT
ILE A T T	LEU TTA	G G T	ILE A T T	ASN A A C	TRP TGGA	ARG C G T G
VAL G T T 380	VAL G T T 9 440	VAL G T A 500	G A A 560	GLY G G G 620	TRP T G G 680	ILE A T T 740
ASN SER VALA A CTCTGT1	ASP GAT	SER A G T	ALA G C T	PHE LYS	PRO ASP SER TRP CCTGATTCTTGG 570 680	SER T C A
ASN A C	VAL GLX 3 T T G G C C 430	ALA G C A	ARG C G C	PHETTC	ASP G A T	GIN C A G
G G T 370	C T T	HIS TYR ALA SER SACTATGCAAGT 490	ASN A A T 550	THR A C T 610	PRO C C T 670	LEU T T G 730
YS ILE LYS GLY ASN SER VAL ILE A A A T C A A G G T A A C T C T G T T A T T 370 380	LY PHE LYS VAL GLY ASP VAL LEU ILE ARG GLU LYS LEU ASN GLU PHE ALA LYS SER VAL 1 GGTTTAAAGTTGGCGATGTTTAATTCGAGAAAATTAAATGAATTTGCCAAAAGTGTAAA 430 440 450 460		EU PRO ASN ASN ARG ALA GLU ILE LEU ILE GLN ILE ASN GLU ASP ASP LYS ALA LYS LEU A TACCAAATAATCGCGCTGAAATTTTAATTCAATGAAGATGATAAAGCAAAATTGG 550 580 600	LA SER LEU THR PHE LYS GLY CATCATTAACTTTCAAGGGGA 610 620	IU LEU GLA PRO ASP SER TRP TRP A A T T A C A A C C T G A T T C T T G G T G G F 90 end truncated GST/D15	LU LYS_ASP LEU GIN SER ILE ARG VASP TYR TYR LEU ASN GLY TYR ALA LYS ALA GIN A GAAAGAAAGCACAAAAAAAAAAAAAAAAAAAAAAAA
ILE A T C	PHETT	YS GLU AAGAG(PRO C C A	SER TCA	LEU TTA dtnu	LYS. A A A
YS. A A	LY G G T	YS A A	EU T A	C A 7	LU A A	LU A G

3/82					
LEU ASN ASP GLU LYS THR LYS VAL ASN VAL THR ILE ASP V CTAAATGAAAAAAAAAAGTTAATGTAACCATTGATG 810 820 830	ASP LEU ARG SER ALA ARG ILE ILE GLY ASN LEU GLY M GACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTA 870 880 890	LEU SER ALA LEU HIS LEU ASN ASP THR PHE ARG ARG S CTTTCAGCATTACATTTAAATGATACTTTCCGCCGTA 930 940 950	ASN ALA ILE LYS ALA LYS LEU GLY GLU ARG GLY TYR GLY S A A T G C A A T T A A A C T T G G A G A A C G C G G T T A C G G T T A C G G T T A C G T T A C G T A 1020	PRO ASP PHE ASP ASP ALA ASN LYS THR LEU ALA ILE THR L CCTGATTTTGATGCAAATAAAACATTAGCGATAACC 1050 1060 1060 1070	ARG LEU THR VAL ARG GIN LEU ARG PHE GLU GLY AGN THR V CGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCG 1110 1120 1130
ILE A T T	GLY GGA	ARG C G C	TYR T A C	ILE A T A	ASN A A T
現 ACC	LEU CTG	PHE TTC(GLY G G T	ALA G C G	GLY GGA
ASN VAL THR NATGTAACCA 830	ASN LEA AATCT 890	ASP THR SATACT1	ARG C G C (LEU T T A 1070	GLU G A A 1130
ASN A A T	GLY G G T	ASP G A T	GW GAA	THR A C A	PHE T T T
VAL GTT	ILE A T A	ASN A A T	GLY G G A	LYS A A A	LEU ARG PHE GLU GLY TTCGCTTTGAAGGAA 20 1130
LYS A A A 820	ILE A T T 880	LEU T T A 940	LEU A C T T 1000	ASN A A T 060	LEU C T T 120
THR ACA	ARG C G C	HIS C A T	LYS A A A 1	ALA G C A	ARG GIN GCCAAC
GLU LYS THR LYS VAL GAAAAACAAAAGTT1 10	SER ALA ARG ILE ILE GLY ASN LEU GLY G T G C A C C C A T T A T A G G T A A T C T G G G A G	SER ALA LEU HIS LEU ASN CAGCATTACATTTAAATG 930	ALA ILE LYS ALA LYS LEU GLY GLU SCAATTAAAGCAAAACTTGGAGAA 990	ASP G A T	ARG C G C
TGAA 810	3 SER TAGT 870	ALA G C A	LYS A A A 30	ASP G A T	VAL GTT(
ASP GAT 81	ARG C G T A 870	SER TCA 93	ILE 1 ATTA 990	PHE T T T 105	THR A C T G 1110
ASN A T C	LEU	LEU LEU TTACTT	ALA G C A	PRO ASP PHE CCTGATTTTC 1050	LEU TTA?
LEU	ASP G A C (ASN AATG		ARG C G T
GIN C A G 800	TYR T A T 860	PRO C C T 920	GAA 980	VAL G T A 1040	ARG C G A 1100
ASP VAL GIN 3ATGTTCAC 800	GIN CAGT	GU GAAC	VAL G T A	SER T C A	G G A
ASP G A T	GLU GLY LEU SAAGGTTTA(CTT	ILE ALA ASP VAL GLU TTGCAGATGTAGA?	ASN A A T	ALA G C T
LYS THR VAAACGO 790	GLY G G T 850	ALA GLU LEU ; C C G A G C T T 910	ALA G C A 970	VAL 3 G T A A 1030	ASP C A T 1090
				ER ALA THR VAL ASN SER VAL GCGCAACGGTAAATTCAGTA 1030 1040	EU VAL VAL ASP ALA GLY ARG TTGTTGATGCTGGACGA 1090
LE THR TTACTA	A A T (SER	ASP GATP	ALA G C A	VAL G T T
T T	AL T A	ET T G 1	ET C	語 G C G	EU TTG

		4	182			
S T 30		L A 20	S A 80	G G	T & OC	s A 50
SN S A T T 1200	NL (r c G 1260	13. 13.	LE S I T A 1380	LE 14	НЕ 1 ГТА 1500	YS 9 N A A 1560
A	S S	D C	A	I &	E	Z Z
YR A T	E C	YS A A	G T	E 50	YR A	£ 5
F E	FA	A	8 9	S	FF	S.
THR TRP TYR ASN S ACTTGGTATAATT 1190	E A	YR A T	- Fi O	AL T A	ως υ	514 7 (C
£ £	⁶ 6	F F	S	> 5	ن ۵	AA
まっ 1961	PHE I T C (1250	AL T A 310	GLU S G A G A 1370	C A 430	GLU 3 A G 1490	ASP 3 A T 1550
A A I	F F F	> ບ [7 C	່າ	. G	່ເອ
7 S	選点	₹ 1	とと	≸ ∪	置し	₹ A
. O	_ E	. D	T A	້ ບ	L A	- €
SLU A	G.	ASP A'	ر و بر	ر و ر	I'YR A	R A O
	. 6	ີ ບ	ິຍ	. 6	T I	~ &
GLN A 80	景っき	WAL TO	17.R A 60	元 こ 20 20	GLY 3 G G T 1480	GLU G A 7 1540
A C 11	THR TACA 1240	A G 13	T T 13	A A 14	G G	T G
GLN GLN PACA1	ARG G	GLU A	GLY TYR GGTTAC 1360	G. G.		PHE T
. U	. E	. E	🕒	_	_ E	C J
ARG G	ASP 3. A	ASP 3 A	E L	T I	ASN A	器产
ຼັບ	A C	→ €-	- Z	ູ້ບ	ູ້ວ	, E
GLN GLU MET ARG GLN GLN GLU GLY ZAGGAAATGCGCCAACAAGAAGGAA 1170 1180	1. T. T. O	24 A 0	မ်ာ် ဗ	EL	M C	₩. C
U A 117	3 C T 1	т 129	же с ттс (1350	N 141	ar v 3 r G 1470	SN V ATG 1530
	A 28 C G	SE A G	1 3 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	AS AA	SS ₹	A A
2 0	E E	×	N O	7. T	≅ 5	× 4
で こ	II A 1	ອ _ອ	A A	G 7 G	A C	ອ _ເ
ARG C G T C	KS A A	ASN GLY SER ASN ASP GLU VAL ASP VAL VAL TYR LYS VAL LA A T G G T A G T G A T G T G G A T G T C G T A T A A A G T G A T G T C G T A T A A A G T G A T G T C G T A T A A A G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T G T C A T C A T G A T G T C A T G A T G T C A T G A T G T C A T G A T	TCAACTTTGGTATTGGTTACGGTACAGAGAGTGGTATTA 1350 1350 1360 1370 1380	A A	7 C	SER LEU GLY GLY ASN VAL PHE PHE GLU ASN TYR ASP ASN SER LYS S GTCTTGGTGGAAATGTTTCTTTGAAACTACGATAACTCTAAAA 1520 1530 1560
ر چ ن	A	A	A	ט	ອ _ບ	_ອ ວ
1150 1160	GLY 3 G A 1220	LE T C 280	SER A G T 1340	YS A A 400	TYR F A T 1460	LEU TTT 1520
	0 0 T	I A I	, s	T A L	ָרָ רָּ	7 5 4
SER THRAGEACTI	Ed F	ည်း	ARG ASN THR GLY GTAACACGGGT? 1330	7. T. T.	SP A T	ម្តី ១
ົ≪	. E	ມິບ	ဗ	ย	ر ت	0, A
SEX G (E A	ASP A	と い	SER C	A 7	T 7
T &	. E	E G	, K	A A	A A	r G
ALA ASP CTGATA 1150	VAL 1 G T 1	11.E	ASN [A A A 1330	A 0 0 0 0	1,YS 3, A, A, 1,450	GLY , G G ' 1510
1 T G) A G 12	; A A 12	13	} ≥ 6 13	G 4	т G 15
ي اي ك	可工	AR G	AR G	OG Y	THE A C	ASE 3 A
SER	Z &	ی ک	о В	۳ ۱	.⊤	. A
വ	GU S	PSA A A	GLI S A	TY	<u>ຕີ</u> ອີ	LYE
AL SER ALA ASP SER THR LEU ARG GLN GLU MET ARG GLN GLN GLU GLY THR TRP TYR ASN S TTTCTGCTGATAGCACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATT 1150 1150 1160 1170 1170	ER GLN LEU VAL GLU LEU GLY LYS ILE ARG LEU ASP ARG THR GLY PHE PHE GLU THR VAL G CACAATTAGTTGAGTTAGGAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTCG 1210 1220 1230 1230	LU ASN ARG ILE ASP PRO ILE ASN GLY SER ASN ASP GLU VAL ASP VAL VAL TYR LYS VAL 1 A A A A C C G A A T T G A T C G T A G T A A T G A T G A G T G G A T G T G	YS GLU ARG ASN THR GLY SER ILE ASN PHE CLY ILE GLY TYR CLY THR GLU SER CLY ILE S A A G A A C G T A A C T A C T T T G G T T A C G G T A C A G A G A G T G G T A T A G A A C G T A G T G G T A T T A C G G T A C A G A G T G G T A T T A C B A C A C A G A G A G T G G T A T T A C B A C A C A C A C A C A C A C A C A	ER TYR GLN ALA SER VAL LYS GLN ASPASN PHE LEO GLY THR GLY ALA ALA VAL SER ILE A GTTATCAAGCAAGTGTTAAACAAGATAATTTCTTGGGAACAGGGGGGGCAGTAAGTA	LA GLY THR LYS ASN ASP TYR GLY THR SER VAL ASN LEU GLY TYR THR GLU PRO TYR PHE T CTGGTACGAAAATGATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTA 1450 1450 1460 1470 1500	HP LYS ASP CLY VAL SER LEU GLY CLY ASN VAL PHE PHE CLU ASN TYR ASP ASN SER LYS (C'TAAAGATGGTGTAAGTCTTGGAAATGTTTTCTTTGAAAACTACGATAACTCTAAAA 1510 1520 1530 1530
A F	ы O		׫	ភ	C	Ħ U
					•	

		5	182	•		
С С 620	SER A A G T A 1680	an G 1 T G 1740	& G 3 A G 1800	S F T 360	A G 320	SN L A C A 1980
所 T T C 16	SER A G	ASN A A '	ARG A G A	GEY G G 3	ARG A G 7	ASN A A (
GLY GGTT	ILE A T T	G G T	ASN A A T	PRO GLY CCAGGT 186	ASP ARG GACAGA 192	GLY G.G.A
THR LEU ACTTTA(1610	LYS A A A	LYS A A A	SER LEU ASN ARG GAGC CTTAATAGAG1790	ILE A T T	PRO LEU CATTA(1910	PHE TTT
THR A C T 1610	ASN A A T 1670	PHE T T T 1730	SER LECAGE CONTRACT	THR A C T 1850	PRO C C A 1910	GLX 3 G T ' 1970
VAL G T T	TYR T A T	LYS A A A	ASN A A C	VAL G T T	TYR T A C	ASN A A T
THR THR TYR GLY SER ASN VALACGACTTACGTATA	GLY HIS THR TYR ASN LYS GGTCATACCTATAAAAA 1660 1670	ASN LEU TYR ILE GLA SER MET LYS PHE LYS GLY ASN GATTTATATATCAATGAAATTTAAAGGTAATG AATTTATATATTCAATGAAATTTAAAGGTAATG 1710 1710 1720 1740	SER PHE GLY TRP ASN TYR ASN FTCTTTTGGTTGGAACTATAACA 1770	ALA SER LEU GLY GLY ARG VAL THR ILE CAAGTCTTGGTGGACGAGTTACTATTC 1830 1850	ALA ASP VAL GIN GLY PHE TYR CAGATGTACAGGGTTTCTACC 1890	GLY TYR ALA ASN GLY PHE GLY ASN LAGGATTTGGAAACA 1970 1960 1980
SER A G T .600	HIS C A T .660	SER T C A 720	ASN 3 A A C 1780	GLY G G A 840	GLY G G T 900	TYR T A T 0
GLY G G A	GLY GGT	GIN CAA 1	TRP TGG	GLY GGT	GLN CAG	GCA 1
TYR T A C	VAL GLY LEU STAGGATTAC 1650	ILE A T T	G G T	LEUCTT	VAL G T A	ALA G C A
THR ACT 0	GLX 3 G A 0	TYR F A T O	RE LTT (SER A G T	ASP 3 A T	SER C T
THR A C G 159	VAL 3 T A (LEU F T A : 1710	SER E FCTT 1770	ALA 3 C A 1 1830	ALA 3 C A 6 1890	ALA SER ALA SCATCTGCA 1950
LYS ARG THR THR TYR GLY SER ASN VAL THR LEU GLY PHE FA A G C C C C C C C C C C C C C C C C C	TYR TYR VAL GLY LEU GLY HIS THR TYR ASN LYS ILE SER P TATTATGTAGGATTAGGTCATACCTATAATAAATTAGTA 1650 1660 1670 1680	P-L	ASP PHE SER PHE GLY TRP ASN TYR ASN SER LEU ASN ARG C GATTTTTCTTTTGGTTGGAACTATAACAGCCTTAATAGAG 1770 1780 1800	VAL LYS ALA SER LEU GLY GLY ARG VAL THR ILE PRO GLY S GTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGTT 1830 1840 1860	LEU SER ALA ASP VAL GIN GLY PHE TYR PRO LEU ASP ARG P CTAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAG 1890 1920	ALA LYS ALA SER ALA GLY TYR ALA ASN GLY PHE GLY ASN I GCAAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACA 1950 1960 1970 1980
LYS A A G	TYR T A T	ARG C G T	ASP G A T	VAL G T T)	LEU	ALA 3 C A 1
ASN TYR AACTAT 1580		ASN A A C 1700		GLY G G G 1820	LYS A A A A 1880	SER T C T (1940
ASN A A C	ASN SER AACTCO 1640	TYR [A T	ASP PHE 3 A C.T.T.1 1760	LYS GLY 1 A A G G (TYR 'A C	VAL SER STATCT 1940
SER T C T A	ASN A A T 1	GLU 3 A A 1	ASN A A T (THR ACTA	TYR FACT	WAL TTG
ER ASP THR SER SER ASN TYR GTGATACATCCTCTAACTAT 1570 1580	RO VAL ASN GLU ASN ASN SER CTGTAAATGAAAATAACTCC 1630 1640	SN PHE ALA LEU GLU TYR ASN ACTTTGCTCTAGAATATAAC 1690 1700	LY ILE LYS THR ASN ASP PHE GCATTAAAACAAATGACTTT 1750 1760	LY TYR PHE PRO THR LYS GLY G C T A T T T C C C A C T A A G G G 1810 1820	ER ASP ASN LYS TYR TYR LYS CTGATAACAAATACTACAAA 1870 1880	SP HIS LEU TRP VAL VAL SER A T C A C C T C T G G G T T G T A T C T 1930 1940
ASP THR	ASN A A T	ALA G C T (LYS A A A A	PHE TTC(ASN A A C 1	LEU CTC1
ASP G A T	VAL G T A	P雅 T T T T	ILE ATTA	LY TYR GCTAT1	ASP 3 A T	HIS A C (
G 5	C T (SN A C	G C Z	C C C	C I	SP A T (

		6/8			
THR TYR THR ALA GLY GLY ILE GLY SER LEU ARG GLY PHE A	ASN ALA ILE TYR ALA GLU TYR GLY ASN GLY SER GLY THR G	LY THR PHE LYS LYS ILE SER SER ASP VAL ILE GLY GLY ASN ALA ILE ALA THR ALA SER A	LA GLU LEU ILE VAL PRO THR PRO PHE VAL SER ASP LYS SER GLN ASN THR VAL ARG THR S	ER LEU PHE VAL ASP ALA ALA SER VAL TRP ASN THR LYS TRP LYS SER ASP LYS ASN GLY L	EU GLU SER ASP VAL LEU LYS ARG LEU PRO ASP TYR GLY LYS SER SER ARG ILE ARG ALA S
CTTATACAGGGGGGCATCGGTTCATTACGTGGTTTG	A A C G C A A T T T A T G G T A A T G G T A G T G G T A C T G	GTACTTTTAAGAAGATAAGTTCTGATGTGATTGGTGGTAATGCAATCGCTACAGCTAGCG	CAGAGTTAATTGTGCAACTCCATTTGTGAGCGATAAGAGCCAAAATACGGTCCGAACCT	CCTTATTTGTTGATGCGCAAGTGTTTGGAATACTAAATGGAAATCAGATAAAATGGAT	TAGAGAGCGATGTATTAAAAGATTGCCTGATTATGGCAAATCAAGCCGTATTCGCGCCT
2010 2020 2030	2070 2080 2100	2110 2120 2130 2130	2170 2180 2180 2190	2230 2240 2250 2250 2260	2390 2330 2340
ARG GLY	G G T	ALA G C T	ARG C G A ,	ASN A A T (ARG C G C G
ARG C G T	SER AGTO	THR ACAG	WAL G T C	ASP LYS 3 A T A A A 1 2270	ARG ILE :GTATTC 2330
LEU	GLY	MA	THR	ASP	ARG
T T A C	G G T	G C T A	A C G	G A T	C G T 7
2030	2090	2150	2210	2270	2330
SER	ASN	ILE	ASN	SER	SER
TCA	A A T	A T C	A A T	T C A (AGCO
GLY	G G T	ALA	GIN	LYS	SER
G G T		G C A A	C A A	A A A	T.C.A.
ALA GLY GLY ILE GLY SER LEU	ILE TYR ALA GLU TYR GLY ASN	ASN	SER	LYS TRP LYS SER	ASP TYR GLY LYS SER SER
SCGGGTGGCATCGGTTCATTAC	TTTTATGCCGAATATGGTAAT	AATG	5 A G C	AAATGGAAATCA(FATTATGGCAAATCAAGC
) 2020	2070	2140	2200	2260	2320
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	GLU GAA	VAL ILE GLY CLY FTGATTGGTGGTA 2130	LYS A A G	LYS A A A	G G C 2
G G T	ALA G C C	G G T	ASP G A T	ASN THR AATACT	TYR T A T
R ALA 2010	E TYR F T T A T 2070	ILE ATT 30	SEX A G C	ASN A A T	ASP G A T 10
THR ACAG 2010	ATT 20	VAL G T G A 2130	VAL (G T G A 2190	TRP 1 T G G A 2250	PRO 7 C C T G 2310
TYR	ALA	ASP	PHE	SER VAL	Leu
TAT!	G C A A	GATO	T T T		TTG
THR	ASN	SER	PRO	SER	ARG
	A A C (T C T (C C A T	A G T	A G A
GLN	PRO	SER	THR	ALA	LYS
C A A	C C T	A G T	A C T (G C A	A A A
2000	2060	2120	2180	2240	2300
PHE TYR GIN TCTATCA 2000	ILE CLY	ILE ATA	PRO C C A	ALA G C G	LEU TTA
PRE	ILE	LYS	VAL	ASP	VAL
TTC		A A G	G T G	G A T	G T A
ARG LEU PRO	TYR GLY SER	THR PHE LYS LYS ILE SER	1LE	VAL	ASP
GTTTACCG1	PATGGTAGTA	ACTTTTAAGAAGATAAGT	1 A T T	G T T	C G A T
1990	2050	2110 2120	2170	2230	2290
LEU FT7	GLY FGG1	PHE.	LEU	PHE TTT	SER
YS ARG LEU PRO PHE TYR GLN	LA TYR GLY SER ILE GLY PRO	THR	LA GLU LEU ILE VAL PRO THR	ER LEU PHE VAL ASP ALA	EU GLU SER ASP VAL LEU
A G C G T T T A C C G T T C T A T C A A A	CTTATGGTAGTATTGGACCTA		CAGAGTTAATTGTGCCAACT	CCTTATTGTTGATGCGG	TAGAGGGATGTATTA?
1990 2000	2050 2060		2170	2230	2290
YS A G C	CTT	LY. G T A	CA	<u>ရ</u> ာ ၁	EU T A

TRP GLN SER PRO ILE GLY PRO LEU VAL PHE SER TYR ALA L	ASN ASP ASP VAL GLUGIN PHE GIN PHE SER ILE GLYGLYS
TGGCAATCTCCTATTGGCCATTGGTATTCTCTTATGCA	AATGATGATGACAGTTCCAATTTAGTATTGGAGGTT
2370 2380 2390 2400	2430 2450 2450 2460
VAL G T A	路口工
LEU	GIN
TTG(CAA1
PRO	PHE
C C A	T T C
380	2440
GLY	GLN
G G G C	C A G 7
ILE ATTG	GLU GAA
PRO	WAL
C C T	GTC(
SER	ASP
TCT	G A T (
237	243)
GLN	ASP
C A A	G A T
TRP	ASN
T G G	A A T
ER THR GLY VAL GLY PHE GLN	GLU
CTACAGGTGTCGGATTCCAA1	G A A 2
2350 2360	2420
PHETTC	TYR TAT
GLY G G A	YS PRO ILE LYS LYS TYR GLU AACCAATTAAAAAATATGAA. 2410 2420
VAL	LYS
' G T C	F A A A ,
2350	2410
GLY	ILE
GGT	ATT/
2	2
THR ACA	S PRO 3
RE O	YS A A (

			7/
	TGCCI	2520	
; ; ;	A S C	0157	
	1 CABAAAA 2500	7007	
TCATCACT	2490		
TTTTTCT	•		
AAATTGAAC			
TTTCTAATA	2470		
ပ			

82

			てききてご
			± ∇ ±
		•	ر - ال
		E E C < C	マーフにつ
		じートマーじ	•
			1
	5 5 5	3 T T C C	
	ر ر ر ر ر ر ر ر ر ر ر ر ر ر ر ر ر ر ر	りりいい	
٠.	A T C A A		
	T C G C 🏲	:	2000
	AAACA		

	G C T T C 2640
	. 1 G C T T C A G G C T A T 2630
じっくしじももるも	2620
	2610
	2600
	2590
•	

G A T C G C 2700	
T C A A C A T C A C C C A G 2690	
АТАТТТТ 2680	
TTAATGCAGGT 2670	-
тт G С т т т С А т 2660	
G A A G A A A A A A A A A A A A A A 2650	
ר פּ ר ד פּ א ר פּ ר די פּ א	

**

붎

臼

8/	82

AAAAGTT	C C A A C A G
2820	2880
A G T A G A A G C	T C A A A A A C G
2810	2870
: T C G T A A A A A A A A 2800	ла G С Т G А Т А Т 2860
GATAAAATTGCTGCTCGTAAAAAGTAGAAGCTAGTT	' G C A C C T T A C G T C A A G C T G A T A T T C A A A A C G C C A A C A G
2780 2780 2800 2800 2810 2820	2840 2870 2880
T A	E.
AAAAGAAGTTG.	C T T T A G A A A A A G 1
2770	2830
AGCA	9999

GATAAAAAA

FIG. 1B

DS-712-2-1 DNA, Eagan D15 sequence IS THE SEQUENCE BEING TRANSLATED

(D. C)	r. -	
TTAACCTTGAAAATATTAGGGAAATTACTTCCTGGCGATTTG 20 30 40 . 50 50	GGGCCAATTTCTATTGCAAAAGGTGCTGGCCCATCAGCAAAT 80 90 100 110	ТТААСТТТАТGGCACTGATTAGTGTAAATTTAGGGATTATG 140 150 160 170 180
<u>-</u>	A	<u> </u>
A 1	ج ن	E E
ပ်	Ö	₹
ن 	A	ဗ
e e	S I	e E
Ē	Æ	¥
၁ ၉	C C 110	T T 170
O ₽	0 0 10	H H
Ē	ဗ	A
C)	r.	A
£ E	ا ت	e.
₽.	Ö	່ບ
A 1	E	F
4 0	00	6 C
0	A G 100	T A 160
5	V	E
· 4 1	A.	&
E	ິວ	E
. [4	ဗ	O ₁
T 7	₽	ຊ ເ
Æ	K	ဗ
A A	. J	T G 150
A	L	A L
5	E-	E
	A I	H
ပ	æ	E
ک د	ບ	9 4
Æ	ΰ	A A
T T 20	S 8	T 40
⊣	E-	7 7
E	ဗ	E
ပ ပ	A A	E
ပ်	E 7	G T G T A T 30
E-	E→	E+
E	T.	ទ
ပ္	A 0	ဗ်ဓ္က
9	E	ب ت ت
7 3	A A	T _
GGACAGCTTTCCCT 10	₹ .	GATTG 13
ල ල	<u>.</u>	9
A	САТТАААТААТТТАА 70	E G
CA	ပ	E
4	F	K

9/82

10/82					
THR A C 420	LEU T T 480	ASP G A 540	HIS C A 600	LYS A A 660	GLY G G 720
THRACG	ASP G A C	ASN A A T	ALA G C G	VAL G T T	ASN A A C O
THR ACA/	GLY 3 G T	ASP 3 A C A	LYS A A A	ASP VAL	ALA 3 C T
THR ACGA)	GIN GLY	THR ACTG	VAL LYS GTGAAA 90	SER CCAG	ASP ATC
GLY 7 3 G T A 410	WAL G TTC 470	VAL 1 TGA 530	ASP V 3 A T G 590	ILE S LTTT 650	LEU P TAG 710
PIE LTCG	ASP GLY VAL HATGGTGTTC	GIN ARG VAL THR ASP ASN ASE CAGCGTGTGACTGACAATGA 530	ASP 7	ILE	ASN LEU ASP AACTTAGATG 710
LEU	SP (ATG	IN A	PHE ASP TCGATO	SER]	LN P A A A
LEU I FTAT 400	VAL A FTGG 460	GLY GG G T C 520	ARG P CGAT 580	LYS SATA	LYS GIN AAACAAI 700
R L	ARG V	ALA G	GLY A	VAL ALA LYS SER ILE ILE SER ASP VAL GTGGCTAAATCGATCATTTCAGATGTT 640 650	LEU LY
S. A	CC	G G	គ្ន	G A	C E
ALA SER GCAAGTÎ	ILE A T T	ARG ALA CGTGCC	SER AGTO	WAL GTG	ALA G C A C
MET LYS LEJ LEJ ILE ALA SER LEJ LEJ PHE GLY THR THR TAGGATACAAAAAACTTCTAATCGCAAGTTTATTCGGTACGACAACG 370 380 390 400	VAL PHE ALA ALA PRO PHE VAL ALA LYS ASP ILE ARG VAL ASP GLY VAL GIN GLY TGTGTTTGCCGCACCTTTTGTGGCAAAGATATTCGTGGATGGTGTTCAAGGT 430 440 450 450	VAL G T T 510	VAL SER GLY ARG PHE ASP ASP VAL LYS ALA HIS GTAAGTGGTCGATTCGATGTGAAAGCGCA 570 580 590 600	VAL 7 T T 330	ILE LYS GLY ASN SER VAL ILE PRO THR GLU ALA LEU LYS GLN ASN LEU ASP ALA ASN A A T C A A G G T A A C T C T G T T G A G C A C T T A A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C A A A C T T A G A T G C T A A C T A A C T A A A C A A A C T T A G A T G C T A A C A A A C T T A A A C T T A A A C T T A A A C T T A A A C T T A A A C T T A A A C T T A A A C A A A C T T A A A A
LEU CTA'A	LYS A A A	PRO C C T (PHE	SER AGCG	THR A C T
LEU	ALA G C A	LEU TTA	LEU PHE TTATTC	WAL	PRO C
LYS A A A C	VAL 3 T G (SER AGTO	SER FCT1	VAL 3 T T C	ILE \ T T (
LYS A A A A 38(PHE L T T G 440	GLU GIN GIN ILE ARG ALA SER LEU PRO AGAACAAATCCGAGCAAGTTTACCT 490 500	VAL ALA ASN ILE VAL ARG SER TGTGGCTAATATTGTCCGCTCT 550 560	LEU (17 C) (20 620	VAL 3 T T A 680
MET A T G	ALA PRO PHE GCACCTTT	ARG	WAL 3 T C (ASP VAL	SER CTCT
ເວິ່ງ	ALA 3 C A (ILE \TCC	ILE VAL	ASP 1 C	ASN SER
370	ALA ; C C C 430	GLU GIN GIN AACAACAAA 490	ASN \ A T A 550	GIN GLU GLY AAGAAGGCG	ILE LYS GLY ATCAAAGGTA 670
T A C	VAL PAE TGTTTG	A A C	ALA ; C T A	IU A A G	YS A A G
≪ .	₩ E	C	æ 5	9	L'A
9	VAL T (GLU A 1	WAL TG(A A	1E 7 C
T A	E G	- 8 8	T G	T C	Y A

ASP GLU LYS THR LYS VAL ASN VAL THR ILE ASP VAL

GLN LEU ASN

THR ASP VAL

LYS

TYR ALA LYS ALA GIN ILE

GLY

ASN

ASI

TYR TYR LEU

ILE ARG ASP

S

LYS ASP LEU GLN

GAAAGATTTGCAGTCAATTCGTGATTATTTAATTAATGGCTATGCCAAAGCACAAAT

		11	182	
LYS A A 780	LEU C T 840			GLU C G A 1020
VAL G T A	THR A C G	LYS LEU ALA AAATTGGC 900	MET A T G	PHE GLU TTCGA 1020
SER AGT	ASN A A T	LYS A A A	GIN C A A	GIN C A A
LYS A A A O	WAL GTC 0	ALA G C A 0	GLU G A A	ALA G C G
ALA G C C	PRO ILE VAL ASN THR LEU: CTATTGTCAATACGCT 830 840	LYS A AAAG 890	GLN C CAAG 950	GLY P G G T G 1010
PHE TTT	PRO C C T	ASP 3 A T	LEU FTA	GLU 3 A A (
GU GAA	GLU G A A	ASP 3 A T	THRACA	PHE I T T (
ASN A A T 760	VAL 3 T T 820	GLU 3 A A (880	SER A G T 9	LYS A A A 9 1000
LEU ILE ARG GLU LYS LEU ASN GLU PHE ALA LYS SER VAL TAATTCGAGAAAATTAAATGAATTTGCCAAAAGTGTAA 750 760 770	ASN ALA THR VAL GLU NACGCAACAGTTGAAC 110 820	ILE LEU ILE GIN ILE ASN GLU ASP ASP LYS ALA TTTTAATTCAATGAAGATGATAAGCAA 870 880 890	ASN GLU SER VAL SER SER THR LEU GLN GLU GLN NACGAATCTGTTAGTAGCAGTACATTACAAGAACAA)	TRP LYS LEU TRP GLY ASN LYS PHE GLU GLY ALA GLN GGAAATTATGGGGAAATAAATTTGAAGGTGCGCAA7 1000 1010
LYS A A A	ALA G C A	ILE A T C	SER A G T	GLY G G A
GLU G A A 750	ASN A A C 810	GLN C A A 870	VAL G T T 930	TRP T G G
ARG C G A	GLY ARC TYR GTCGCTATA	ILE A T T	SER T C T	LEU TTA
ILE A T T	ARG C G C	LEU TTA	GLU GAA	LYS A A A
LEU TTA 10	G G T	ILE ATT 50	ASN A A C	TRP TGG
VAL G T T	VAL G T A G 800	GLU GAA 86	G G G A 920	TRP 7.
ASP GAT(SER AGT(ALA G C T (LYS A A G (SER
0 2 9 9	ALA G C A A	ASN ARG AATCGCG 850	PHE TTC?	ASP GAT1
VAL G T T (730	HIS TYR	ASN A A T 850	SER LEU THR CATTAACTT	PRO C C T C 970
LYS A A A C	HIS C A C	ASN AATA	LEU TTA	GIN CAA(
PHE LYS VAL GLY ASP VAL LEU IIE ARG GLU LYS LEU ASN GLU PHE ALA LYS SER VAL LYG GTTTAAAGTTGGCGATGTTTAATTCGAGAAAATTAAATGAATTGCCAAAAGTGTAAA 730 770 780	GLU HIS TYR ALA SER VAL GLY ARG TYR ASN ALA THR VAL GLU PRO ILE VAL ASN THR LE A G A G C A C T A T G G T C G C T A T A A C G C A A G T T G A A C C T A T G T C A A T A C G C T 790 830 840	PRO ASN ASN ARG ALA GLU ILE LEU ILE GLN ILE ASN GLU ASP ASP LYS ALA LYS LEU ALA A C C A A A T A A T T T T A A T T C A A T G A A G A T G A T A A G C A A A T T G G C 800 850 850 860 860 870	SER LEU THR PHE LYS GLY ASN GLU SER VAL SER SER SHR LEU GLN GLU GLN MET GLA ATCATTAACTTTCAAGGGGAACGAATCTGTTAGTAGCAGTACATTACAAGAACAAATGGA 910 920 930 930	LEU GIN PRO ASP SER TRP TRP LYS LEU TRP GLY ASN LYS PHE GLU GLY ALA GLN PHE GLA ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATAAAT

		12/	82		
MET A T 1200	SER 7 A G 1260	SER F A G 1320	LEU C T 380	VAL 3 G T 1440	SER 1 T C 1500
GLY 3 G T	ARG C G T	GLY 3 G T	THR ICC	選() () () () () () () () () () () () () () (ASN SER ATTC 1500
GLY G A (ARG G C (ARG CLY TYR CLY GCGGTTACGG1	ALA ILE THR LEU SCGATAACCCT)	ARG PHE GLU GLY ASN THR VAL GCTTTGAAGGAAATACCGT 1430 1440	TR A T A
EU T G G	PHE ARG	LY G T T	.∵ ≰ 8 : 10 :	LY Z	W 1
N L L I C 1190	IR P	ें इट द 1310	U AI AG (1370	JJ GI NAG (1430	R TF C T T T T T T T T T T T T T T T T T
/ AS TAA	ASN ASP THR NATGATACT1	- AR A C G	LYS THR LEU AAACATTAG 1370	T G A	THE A A C
. GL)	ASE	GLY GLU	THR	PHE T T	G G 7
ILE A T A	ASN A A T	GLY GGA	LYS A A A	ARG C G C	GLU G A A
ILE A T T 1180	LEU F T A A 1240	LEU C T T 1300	ASN A A T 1360	LEU 7 T T 1420	GIN GLU GLY THR TRP TYR PAAGGAACTTGGTAT 1480 1490
SP LEU ARG SER ALA ARG ILE ILE GLY ASN LEU GLY GLY MET A C C T T C G T A C G C A T T A T A G G T A A T C T G G G A G G T A T A T A G G T A A T C T G G G A G G T A T 1190 1170 1200	LEU HIS PTACATI	ASN ALA ILE LYS ALA LYS LEU NATGCAATTAAAGCAAACTT 1290	PHE ASP ASP ALA ASN TTGATGATGCAAAT 1350 1360	SIN A A C	GIN C A A C
LA CAC	EU TAC	LA CAA	SP A T G	ည်း	ARG C
SER 7	A T C	KS P AAG	SP A	L A	E B
ARG S CTA (SER ALA ICAGCAT 1230	3 LYS T A A 1290	3 AS T G 7	T G T 1410	' MET A A T 1470
T C G	SE TTC	ILI	E L L L	THE L	GLU G A
ASP LEU SACCT 1	LEU LEU PRACTT	ALA G C 7	ASP FGATT	Led TTP	GLN C A G
ASP GAC	LEU TTA	ASN A A T	PRO C C T 0	ARG C G T 0	ARG C G T
TYR P TATG 1160	PRO 1 C C T T 1220	GLU P 3 A A A 1280	VAL 3 T A 134	ARG ARG LEU THR VAL ARG GLN GACGTTTAACTGTTCGCCAA 1400	LEU ARG GIN GLU MET TTACGTCAGGAAATG 1460 1470
ASN CLU CLY LEU CLN TYR A AAATGAAGGTTTACAGTATG1 1150 1160	SFR ALA GLU LEU GLU PRO LEU LEU SER ALA LEU HIS LEU ASN ASP THR PHE ARG ARG SEI GTCTGCCGAGCTTGAACCTTTCAGCATTACATTAAATGATACTTTCCGCCGTAG 1210 1220 1230 1230	ASP ILE ALA ASP VAL GLU ASN ALA ILE LYS ALA LYS LEU GLY GLU ARG GLY TYR GLY SER TGATATTGCAGATGTAGAAAATGCAATTAAAGCAAACTTGGAGAACGGGTTACGGTAG 1270 1220 1320	ALA THR VAL ASN SER VAL PRO ASP PHE ASP ASP ALA ASN LYS THR LEU ALA ILE THR LEI CGCAACGGTAAATTCAGTACCTGATTTTGATGATGCAAATAAACATTAGCGATAACCT 1330 1330 1340 1350 1380	VAL VAL ASP ALA GLY ARG ARG LEU THR VAL ARG GLN LEU ARG PHE GLU GLY ASN THR VAL TGTTGTTGATGCTGGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGT 1390 1400 1400 1410	SER ALA ASP SER THR LEU ARG GIN GLU MET ARG GIN GLU GLY THR TRP TYR ASN SET TTCTGCTGATAGCACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATTC 1450 1450 1460 1470 1500
LEU TA(LEU TTG	ASP A T G	A T T	LA CTG	G C A
GLY 5 G T T 1150	LU 1 A G C 210	ALA 7 3 C A G 1270	AL 7 F A A 330	SP A T G 390	SP S V T A I50
CU CO	A G	E 1 G 1	R V G G .	L AS TG1	ALA AS SCTGA
5 7 E	SER ALA GLU CTGCCGAGO	ASP ILE ATATT	ALA THR VAL CAACGGTAA 1330	VAL VAL TTGTT(AL r G C
AST 1 A A	SEA 3 T C	ASF	ALA G C	VAL G T'	SER T C T (
~	9	F.		E-	₽

		13/8			
GLU GA 1560	LYS S A A 1620	SER [A G 1680	ALA \ G C 1740	THR 7 A C 1800	SER A G 1860
VAL G T C	VAL G T C	ILE A T T	ILE A T A	PHE ITT	LYS A A A
THR A C A	LYS A A A	SER GLY ILE AGTGGTATTA	SER A G T	PRO TYR PHE CCTATTT1	SER
PHE GLU THR VAL GLU TCGAAACAGTCGA 1550 1560	TYR TAT 0	SER A G T 0	ALA VAL SER SCAGTAAGTA 1730	04d 0 0 0 0	ASN A A C '
'XS ILE ARG LEU ASP ARG THR GLY PHE PHE GLU THR VAL GLA A A A T T C G C T T A G A T C G T A C A G G T T T C T T C G A A C A G T C G A A C A G T C G A A C A G T C G A	SN GLY SER ASN ASP GLU VAL ASP VAL VAL TYR LYS VAL LYS A T G G T A G T G A A G T G G A T G T C G T A T A A A G T C A A 1590 1620	ILE GLY TYR CLY THR CLU S TTGGTTACGGTACAGAGA 1660 1670	ALA V. G C A G 1730	GLY TYR THR GLU F GGTTATACCGAGC 1780 1790	X GLY ASN VAL PHE CHU ASN TYR ASP ASN SER LYS SEF TGGAAATGTTTTCTTTGAAACTACGATAACTCTAAAG 1830 1860
PHE TTCT	VAL GTC	THR A C A	THR GLY ALA ACAGGGCGG 1720	THR ACC	TYR T A C
ARG THR GLY : G T A C A G G T 5	ASP GAT	G G T	aty G G G	TYR T A T	ASN A A C
THR A C A 1540	VAL G T G 1600	TYR T A C 1660	THR A C A 1720	G G T 1780	GLU G A A 1840
ARG C G T	ASP GLU VAL ATGAAGTGO	G G T	GLY G G A	ASN LEU AATTTG	PHETT
ASP GATC	ASP G A T	ILE A T T	PHE LEU GLY FTCTTGGGAA 710	ASN A A T	PHE T T C
ARG LEU : G C T T A C 1530	ASN A A T (1590	GLY F G G T 1650	PHE F T T C 1710	SER VAL GTGTCA 1770	VAL F G T T 1 1830
ARG C G C	SER AGT 1	PAE エエエ 16	ASN A A T	SER A G T	ASN A A T
LYS ILE AAAATTC	C C T	ASN AACT	ASP G A T	THR A C G	GLY F G G A
LYS A A A A 20	ASN A A T 80	ILE ATC 10	GLN C A A	G G T	G G T
VAL GLU LEU GLY 1 3 T T G A G T T A G G A A 1510 1520	ASP PRO ILE ASN GLY FATCCTATCAATGG7 1580	SER] A G T A 1640	ALA SER VAL LYS GLN ASP ASN SCAAGTGTTAAACAAGATAAT1 1690 1700 1700 17	ASN ASP TYR GLY THR AATGATTATGGTACG 1760	LEU (CTTG 1820
TTA	PRO CCT	ASN THR GLY AACACGGGT 1630	VAL G T T	ASP G A T	SER A G T
GLU GAG	ASP G A T	THR ACG	SER A G T	ASN A A T	VAL G T A
VAL G T T 1510	ARG ILE GAATT (1570	ASN A A C 1630	ALA G C A 1690	LYS A A A 1750	GLY G G T 1810
LEU TTAG	ARG C G A	ARG C G T 1	GLN CAAC	GLY THR LYS G T A C G A A A I 1750	LYS ASP GLY VAL SER LEU GLY AAGATGGTGTAAGTCTTGG1 1810 1820
GLN LEU VAL GLU LEU GLY L ACAATTAGTTGAGTTAGGAA1 1510 1520	ASN ARG ILE ASP PRO ILE AS AAACGAATTGATCCTATCA1 1570 1580	GLU ARG ASN THR GLY SER ILE ASN PHE GLY ILE GLY TYR CLY THR GLU SER GLY ILE SER A G A A C G T A A C G G G T A T C A A C T T T G G T A T T G G T T A C G G T A C A G A G T G G T A T T A G 1630 1640 1680	TYR GLN ALA SER VAL LYS GLN ASP ASN PHE LEU GLY THR GLY ALA ALA VAL SER ILE ALA TTATCAAGCAAGTGTTAAACAAGATAATTTCTTGGGAACAGGGGGGGG	GLY THR LYS ASN ASP TYR GLY THR SER VAL ASN LEU GLY TYR THR GLU PRO TYR PHE THE TGGTACGAAAAATGATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTAC 1750 1750 1760 1770 1770	LYS ASP GLY VAL SER LEU GL TAAAGATGGTGTAAGTCTTGG 1810 1820

GLY TYR ALA ASN GLY PHE GLY ASN LYS

TCACCTCTGGGTTGTATCTGCAAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAA

SER ALA LYS ALA SER ALA

TRP VAL VAL

E C

TGATAACAAATACTACAAACTAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGA

ASP

TYR PRO LEU ASP ARG

閚

GIN

M.

ALA ASP

SEX

ASP ASN LYS TYR TYR LYS LEU

FIG.1B.(CONTINUED)

		14/	82	
PRO	ASN	GLY	GLY	SER
C C C	2 A A	r G G	G G	'T C
1920	1980	2040	2100	2160
雅 LTC 1	SER AGT	ASN A A T	ARG A G A	GLY G G T
GLY	ILE	GLY	ASN	PRO
3 G T		3 G T	A A T	C C A
LEU TA	LYS	LYS A A A (LEU	ILE ATT(
THR LEU ACTTTA 1910	ASN E) A T A 1	PHE I TTA 2030	SER LI GCC 2090	THR I
VAL	TYR	LYS	ASN	VAL
TTB	'ATA	. A A T	A C A	; T T ?
ASN	TR	MET	TYR	ARG
A T G	CCT	TGA	'ATA	: G A G
SER	HIS	SER	ASN TYR ASN SER LEU ASN ARG GLY	GLY
A G T A	2 A T A	C C A A	A A C T A T A A C A G C C T T A A T A G A G G	3 G A C
1900	1960	2020	2080 2080 2100	2140
SER ASN TYR LYS ARG THR THR TYR GLY SER ASN VAL THR LEU GLY PHE PROPERATETA CTATA GGTTTCCC CTAACTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCCC 1880 1890 1890 1900	ASN ASN SER TYR TYR VAL GLY LEU GLY HIS THR TYR ASN LYS ILE SER ASN A A T A A C T C C T A T T A G G A T T A G G T C A T A C C T A T A A A A T T A G T A A 1980 1940 1950 1980	RG ASN LEU TYR ILE GLN SER MET LYS PHE LYS GLY ASN GL) GTAATTTATATATTCAATGAAATTTAAAGGTAATGG 2010 2010 2020	ASN ASP PHE ASP PHE SER PHE GLY TRP ASN TYR ASN SER LEU ASN ARG GLA A A T G A C T T T G A T T T T C T T T T G G A C T A T A A C A G C C T T A A T A G A G G 2060 2070 2060	GLY 3 G T Q
TYR	LEU	ILE \ T T C	GLY 3 G T 1	LEU
THR	GLY	TYR	PHE	SER
3 A C T 1	A G G A 1	1 T A T 1	LTTT(A A G T (
1890	1950	2010	2070	2130
THR	VAL	LEU	SER	ALA
ACG1	3 T A (LTA'	ICT	G C A 2
ARG	TYR	ASN	PRETT	LYS
C G T 2	I A I (A A T		A A A (
LYS	TYR	ARG	ASP	VAL
A A G	TAT	C G T	G A T	G T T
TYR I T A T A 1880	SER 7 T C C T 1940	ASN AACC 2000	PHE 7 TTTG 2060	GEGG 2120
ASN	AGN	TYR	ASP	LYS
A A C	A A C	T A T	G A C	A A A
SER	ASN	GLU	ASN	THR
	A A T	GAA	A A T	ACT
SER T C C 1	ASN GLU AATGAA?	LEU C T A 1990	THR A C A 2050	PRO C C A 2110
ASP THR SER ATACATCC1 1870	ASN A A T	PHE ALA LEU GLU TYR ASN ARG ASN LEU TYR ILE GLN SER MET LYS PHE LYS GLY ASN GLY TTGCTCTAGAATTTAATTTATATTCAATCAATGAAATTTAAGGTAATGG 1990 2000 2010 2010	ILE LYS THR TTAAAACAA 2050	TYR PHE PRO THR LYS GLY VAL LYS ALA SER LEU GLY GLY ARG VAL THR ILE PRO GLY SER ATTTCCCAACTAAAGGGGTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGTTC 2110 2120 2130 2130
ASP THR SER SER ASN TYR LY	VAL ASN GLU ASN ASN SER T	PHE ALA LEU GLU TYR ASN AF	ILE LYS THR ASN ASP PHE A	TYR PHE PRO THR LYS GLY VAL LYS ALA SER LEU GLY GLY ARG VAL THR ILE PRO GLY SE
TGATACATCCTCTAACTATAI	TGTAAATGAAAATAACTCCTI	CTTTGCTCTAGAATATAACCC	CATTAAAACAAATGACTTTG	CTATTTCCCAACTAAAGGGGTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGTTC
1870 1880	1930 1940	1990 2000	2050 2060	2110 2120 2130 2130

SUBSTITUTE SHEET

		15/8			
ALA	GLY	ALA	SER	LEU	SER
F G C	r G G	: G C	2 T C	A T T	T C
2340	2400	2460	2520	2580	2640
PHE TTTC 23	THR ACT(SER AGCO	THR ACC	GLY G G A	ALA G C C
ARG GLY GTGGTT	SER GLY	THR ALA CAGCT	ARG C G A	ASN A A T	ARG ALA
ARG	SER	THR	VAL ARG THR	LYS	ILE
C.G.T.	A G T	ACA	STCCGAACC		TT(
HR TYR THR ALA GLY GLY ILE GLY SER LEU ARG GLY PHE ALA	TYR GLY ASN GLY SF	ALA 1	GIN ASN THR V	R VAL TRP ASN THR LYS TRP LYS SER ASP LYS ASN GLY LEU	ARG ILE
CTTATACAGCGGGGGATCGGTTCATTACGTGGTTTGC	ATGGTAATGGTAC	G C T A	SAAAATACGG	TGTTTGGAATACTAAATGGAAATCAGATAAAATGGATT	CGTATTC
2310 2320 2330	2380	2450	2510	2550 2580	2630
SER	ASN	AIA ILE	ASN	SER	SER
TCA	A A T		A A T	T C A	A G C
GLY G G T T	GLY G G T	ALA G C A	GLN C A A	LYS A A A	LYS SER SER AAATCAAGCO 2620
ILE	TYR	ASN	SER	TRP	LYS
A T C G	T A T	A A T C	A G C (T G G A	A A A 2
2320	2380	2440	2500	2560	2620
CLY CLY	ALA GLU	G G T	ASP LYS	LYS A A A	GLY 3 G C /
G G T	ALA G C C	ILE GLY GLY \TTGGTGGT!	ASP G A T	THR ACTI	ASP TYR GLY SATTATGGCA S10
ALA	TYR	ILE	SER	TRP ASN	ASP
1 G C G C	F T A T (3 A T T	3 A G C 6	I G G A A T	7 G A T
2310	2370	2430	2490	2550	2610
THR ACA(ILE	WAL GTGA 24	WAL GTG	TRP T G G	PRO C C T G 26
TYR	AIA	ASP	PHE	VAL	LEU
FTAT 2	G C A A	GATG	T T T	GTT	TTG
THR ACT	ASN A A C	SER .	PRO C C A	SER A G T 0	ARG A G A O
PHE TYR GIN THR	PRO 7	ILE SER S	PRO THR PRO PHE VAL	ALA S	ASP VAL LEU LYS ARG
TCTATCAAACT	C C T A	LTAAGTT	CAACTCCATTTGTGA	G C A A	BATGTATTAAAAGA
2300	2360	2420	2480 24	2540	2590 2600
TYR	ILE GLY	ILE	PRO	ALA	LEU
T A T	TTGGA	A T A	C C A	G C G	TTA
PHE	ILE	LYS	VAL	ASP	VAL
TTC	A T T	A A G A	GTGC	3 A T	3 T A
PRO	SER	LYS	11.E	VAL	ASP
C C G 1	A G T A	A A G <i>I</i>	A T T G	3 T T (3 A T (
2290	2350	2410	2470	2530	2590
LEU	G G T	PEE TTT	LEU FTAA	LEU PHE VAL ASP ALA ALA SER TATTGTTGATGCGGCAAGT 2530 2540	SER AGCG
ARG LEU PRO PHE TYR GIN TO G C G T T T A C C G T T C T A T C A A A (2290	TYR GLY SER ILE GLY PRO ASN ALA ILE TYR ALA GLU TYR GLY ASN GLY SER GLY THR GL) TTATGGTAGTATTGGACCTAACGCAATTTATGCCGAATATGGTAATGGTAGTGGTACTGG 2350 2350 2360 2370 2370	THR PHE LYS LYS ILE SER SER ASP VAL ILE CLY CLY ASN ALA ILE ALA THR ALA SER ALA TACTTTTAAGAAGATAAGTTCTGATGTGATTGGTGGTAATGCAATCGCTACAGCTAGCGC 2410 2420 2430 2430 2460	GLU LEU ILE VAL PRO THR PRO PHE VAL SER ASP LYS SER GLN ASN THR VAL ARG THR SET A G A G T T A A T T G T C C A T T T G T G A G C G A T A G A G C C A A A A T A C G G T C C G A A C C T C 2470 2480 2520	LEU PHE VAL ASP ALA ALA SE CTTATTGTTGATGCGGCAAG 2530 2540	GLU SER ASP VAL LEU LYS ARG LEU PRO ASP TYR GLY LYS SER SER ARG ILE ARG ALA SE A G A G C G A T G T A T T A A A A G A T T G C C T G A T T A T G G C G A A A T C A A G C C G T A T T C G C G C C T C 2590 2690 2600 2610 2610
			-	-	•

Ш
\supseteq
\leq
E
Z
Ö
\mathcal{O}
B.(
<u>G</u>
正

LYS	A A 2700	SER ' T C 2760
ALA	TGGCAATCTCCTATTGGGCCATTGGTATTCTCTTATGCCAA	GLY G G T
TYR	T A T	GLY GGA
SER	T C T	11.E A T T 50
器	ттст 2690	SER A G T A 2750
VAL	GTA	PHE TTT
LEU	T T G	GIN C A A
PRO	C C A 2680	PHE T T C 2740
GLY	9 9 9	GLN C A G
ILE	ATT	GW GAA
PRO	ССТ 2670	VAL G T C 2730
SER	T C T	ASP G A T
GLN	CAA	ASP GAT
TRP	ក 08 ខ	ASN A A T
GIN	C A A 266	GU GAA 277
꿆	T T C	TYR T A T
GLY	G G A	LYS A A A
VAL	G T C 2650	LYS A A A 2710
GLY	G G T	ILE A T T
THR	TACAGGTGTCGGATTCCAA 2650 26	PRO ILE LYS LYS TYR GLA ACCAATTAAAAAATATGA 2710

1	6/82	
GCCTAAT	2820	AATGAAA
TTCTTCATCAGAACTCAAAAACAACGTTCTCTGCCTAAT	2810	TAAACCCATCATTTAATTAAGGATATTTATCAAATGAA
AACTCAAAA	2800	TTTAATTAA
TCTTCATCAG	2790	AAACCCATCA
TTGAACTTTT	2780	GAGAAAATATI
TTTCTAATAAA	2770	TTAATTGGGCA

2870

2840

ပ	2940
TT	
TGC	
TAT	
GCJ	0
G	2930
TCA	
GCTT	
Ħ	
ACT	2920
1 2 5	23
T T	
TA	
S	_
TTA	2910
T G	
CT	
GCAC	906
ပ ပ	~;
A A	
G T	
AAA	_
CAZ	2890
ဗ	
AT	
A C	
$ \leftarrow $	

SUBSTITUTE SHEET

图

FIG. 1C

DS-691-1-5 DNA, Minn A D15 sequence IS THE SEQUENCE BEING TRANSLATED

		17/8	2	•		
GCAAAA	09	C A C T G A T T A 120	G T C A T T T A G 180	A A A G C A T C T 240	TTAATGATT	300
GTGGGCCAATTTCTATT	20	TTGGATTGGTATTTTTTAAGTTTTATGGCA 100 110	TTAGATGGCGGT	G A G C G G G T G C A A A 230	GTTTGCATTATTAA	290
AATAATTTAAGTGG		G T A T T T T T A 100	C A T T A C C A G T A T T A 160	ААССТGТТТСТGАG 220	GCTTAACGGTG	280
CATTA			A T T T A T T T C 150	G T T A A A G G A A A 210	TGTTATTAA	270
GGCGATTI		G C A C A T C A G C A A A T A 70 80	AGGGATTATGA 140		GGCGCAGCAC	260
⊢		G T G C T G G C A C A C A C A C A C A C A C A C A	GTGTAAATTTA 130	тттттаасаат G G A A G C T 190 200	GTTATCGAATTGGCG	250

		18/82			
LEU TTA 360	VAL G T G 420	GLY G G T 480	ARG C G A 540	LYS A A A 600	LYS A A A 660
SER AGT	ARG C G T	ALA G C C	GLY GGT	ALA LYS GCTAAA 600	LEU CTT
ALA G C A A	ILE a t t	ARG C G T	SER AGT	VAL G T G	ALA G C A
MET LYS LEO LEO ILE ALA SER TGAAAAACTTCTAATCGCAAGT 340 350	LYS ASP AAAGAT 410	GIN GIN IIE ARG ALA SER LEU PRO VAL ARG ALA GLY CAACAAATCCGAGCAAGTTTACCTGTTCGTGCCGGT 450 460 470	LEÙ PHE VAL SER TTATTCGTAAGT 530	LEU VAL VAL SER VAL CTTGTTGTTAGCGTT 580 590	SILE LYS GLY ASN SER VAL ILE PRO THR GLU ALA LEU LYS A A T C A A G G T A A C T C T G T T A T T C C C A C T G A G C A C T T A A A 640 650 660
CTA	LYS A A A	PRO C C T	PAE TTC(SER AGC	THRACT
	ALA 3 C A	LEU TTA	LEÙ TTA1	VAL G T T	PRO C C C A
LYS A A A O	VAL G T G O	SER AGT	SER TCT'	VAL G T T	ILE ATTC
MET LYS LYS LEU TAGGATACAATGAAAAACTT 330 340	PHE VAL TTTGTG	ALA S G C A A 460	ARG SER CGCTCT 520	LEU V CTTG 580	VAL : G T T A 640
MET A T G	PRO C C T	ARG C G A	WAL 3 T C		SER TCT
r c G 7	ALA G C A (ILE	ILE A T T (ASP G A T	ASN A A C
C A A 7	ALA G C C (390	GLN C A A .	ASN A A T 510	GLY G G C 570	GLY G G T 630
ATA	PHETTT	GLN C A A	ALA 3 C T	GLU GLY ASP VAL GAAGGCGATGTG 570	LYS A A A
A G G	VAL G T G	I GLU AGAA	VAL G T G	GIN TCAA	ILE A T C A
			_	HIS C A T 560	LYS A A A 620
TTA	THR ACG	ASP G A C	ASP ASN BACAATO	ALA G C G C	ASP VAL
AAT	THR ACAA	G G T G	ASP G A C	LYS A A A G	ASP G A T
T A T 0	THR ACG1	GLN C A A C	VAL THR 3 T G A C T C 490	VAL GTG1	SER TCAG
G T C T 310	GLY 1 G G T A 370	VAL (GTTC	VAL GTG 49	ASP V G A T G 550	ILE : ATTT 610
T A C	PETCO	G G T (GIN ARG VAL THR ASP ASN ASI CAGCGTGTGACTGACAATGA 490 500	PHE ASP ASP VAL LYS ALA HES TTCGATGATGTGAAAGCGCA 550 560	SER ILE ILE SER ASP VAL LYS TCGATCATTTCAGATGTTAA 610 620
T T T T A C G T C T A T A A T T T A T A S 20	LEU PHE GLY THR THR THR THR TTATTCGGTACGACAACGACT 370 380	ASP GLY VAL GLN GLY ASP LEI GATGGTGTTCAAGGTGACTT 430	GIN C A G (PHE TTC	SER T C G

		19/82			
ASN A A T 720	VAL G T T 780	GLU G A A 840	SER A G T 900	LYS A A A 960	ASN A A T 1020
LEU TTA	THR ACA	AGN A A T	SER A G C	ASN A A T	ASN A A T
LYS A A A	ALA THR GCAACA	ILE A T C	SER AGT	G G A	LEU TTA
LYS VAL GLY ASP VAL LEU ILE ARG GLU LYS LEU ASN A A A G T T G C G A T G T T T T A A T T C G A G A A A A T T A A A T 690 720	ASN A A C 770	LEU ILE GIN ILE ASN GLU TTAATTCAAATCAATGAA 830 840	LYS GLY ASN GLU SER VAL SER SER A A G G G A A C C G A T C T G T T A G T A G C A G T 880 890 900	TRP T G G 950	ASP LEU GLA SER ILE ARG ASP TYR TYR LEU ASN ASN GATTTGCAGTCAATTCGTGATTATTATTAAATAAT 990 1020
ARG C G A (TYR TAT	ILE A T T	SER T C T	LEU TTA	TYR TAT
ILE A T T	ARG C G C	LEU TTA	GU GAA	LYS A A A	ASP G A T
. LEU TTTAA 700	, GLY AGGT 760	1 ILE A A T T 820	G A A C 880	TRP GTGGA 940	ARG C G T
VAL G T T 1 700	VAL G T	GLU GAAA 820	G G G	TR T G	11.E 7 ATTC 1000
ASP G A T	SER A G T	ASN ASN ARG ALA GLU ILE A A T A A T C G C C C C G A A A T T 810 820	LYS A A G	SER T C T	SER TCA/
G G C G	ALA G C A	ARG C G C	PRE TTC	ASP G A T	GLN C A G
VAL G T T (TYR T A T 750	ASN A A T (810	THR A C T 870	GIN PRO CAACCT 930	LEU T T G (
LYS A A A	HIS CAC	ASN A A T	LA SER LEU THR PHE CATCATTAACTTTC 0 870		ASP G A T
PHETT	YS GLU A A G A G O	PRO C C A	SER TCA	LU LEU AATTA O	U LYS AGAAA O
ASN GLY ACGGG 680	LYS A A A 740	LEU C T A 800	ALA G C A 860	GAA 920	G. A. G. 980
ASN A A C	SER VAL	THR ACG	ALA LYS LEU SCAAAATTG(MET A T G	PHETTC
ALA G C T A	SER A G T	ASN A A T	LYS A A A	GLU GIN SAACAA!	GLN C A A
1 ASP A G A T G 670	A LYS C A A A A 730	PRO ILE VAL ASN THR CTATTGTCAATACGO 790	3 ALA A G C A 850	и спи А G A A 910	r ALA TGCG 970
LEU TTAC	ALA G C C 1	ILE ATT	LYS A A A 8'	GIN C A A C 910	GLY G G T 9
GIN ASN LEU ASP ALA ASN GLY PHE CAAAACTTAGATGCTAACGGGTTT 670 680	GLU PHE ALA LYS SER VAL LYGAAAGTGTAA1 GAATTTGCCAAAAGTGTAA1 730		ASP ASP LYS ALA LYS LEU AI GATGATAAAGCAAAATTGG 850	THR LEUGINGINMETG ACATTACAAGAACAAATGG1 910	PHE GLU GLY ALA GLN PHE GLU TTTGAAGGTGCGCAATTCGAG 970 980
GIN C A A A	G A A	GU GU	ASP G A T	THR ACAT	PHETT

		20/82	•		
LYS A A A 1080	ILE A T T 1140	LEU T T A 1200	LEU CTT 1260	ASN A A T 1320	LEU C'T T 1380
THR A C A	ARG ILE CGCATT 1140	HIS CAT	LYS A A A	ALA G C A	GLN C A A
THR LYS THR ASP VAL GIN LEU ASN ASP GLU LYS THR LYS ACTAAAACGGATGTTCAGCTAAATGATGAAAAAAAAAAA	ARG SER ALA CGTAGTGCAC	SER ALA GLU LEU GLU PRO LEU LEU SER ALA LEU HIS LEU TCTGCCGAGCTTGAACCTTTACTTTCAGCATTACATTA 1170 1180 1190	ALA G C A A	ALA THR VAL ASN SER VAL PRO ASP PHE ASP ALA ASN GCAACGGTAAATTCAGTACCTGATTTTGATGATGCAAAT 1290 1300 1300	VAL VAL ASP ALA GLX ARG ARG LEU THR VAL ARG GIN LEU TGTTGTTGATGCTGGACGTTTAACTGTTCGCCAACTT 1350 1360 1360
GLU 1 G A A A 1070	SER TAGT(ALA 1 G C A T 1190	ILE LYS ATTAAA(PHE ASP [TTGAT(LEU THR VAL TAACTGTT 1370
ASP GATG	ARG C G T	SER TCAG	ILE ATT	PHE TTT	THR ACT
ASN A T G	LEU CTT(LEU	ALA G C A	ASP GATT	LEU TTA
LEU CTA	ASP GACC 0	LEU TTA	ASN AAT	PRO C C T	ARG CGTT
VAL GIN LEU STTCAGCTAI	GLU GLY LEU GIN TYR ASP LEU GAAGGTTTACAGTATGACCTT 1110	GLU PRO LEU FGAACCTTTAC 1180	GLU ASN GAAAAT 1240	SER VAL PRO CAGTACCT 1300	ARG 7 C G A C 1360
VAL G T T	GLN C A G	GLU GAA	VAL G T A	SER T C A	GLY G G A
ASP GATG	LEU TTA	LEU	ASP G A T (ASN A A T 1	ALA G C T
THR LYS THR ACTAAAACGO	GLY G G T 1110	GU LEU GAGCT7 1170	ALA G C A (VAL G T A 7	ASP G A T 1350
LYS A A A	GLU GAA	ALA G C C	ILE A T T	THR ACG	VAL G T T
ACT	ASN A A T	SER TCT(ASP G A T	AIA G C A I	VAL G T T
E⊣	VAL r G T A 1100	MET 'A T G 1160	SER F A G T 1220	SER 7 A G C 1280	LEU ; C T T 1340
ALA GIN ILE SCACAAAT 1040	ASP VAL G A T G T 7	GLY GLY MET 3 G A G G T A T (ARG C G T	TYR CLY SER FACGGTAG 1280	ILE THR LEU NTAACCCT 1340
ALA G C A	ILE A T T C	G G A	ARG CGCC	TYR T A C	ILE A T A
TYR ALA LYS PATGCCAAAG 1030	VAL THR GTAACC 1090	GLY ASN LEU GTAATCTGO 1150	ASP THR PHE (ATAC)	GLU ARG CLY 3 A A C G C G G T T 1270	ALA G C G
ALA I G C C A 1030	VAL 1 G T A A 1090	ASN 1 AATC 1150	THR FACTT 1210	ARG C C G C G 1270	LEU 1 TTAG 1330
	ASN A A T		rn		LYS THR LEU ALA ILE THR LEU AAAACATTAGCGATAACCCT 1330 1340
G G C 1	VAL G T T	ILE A T A C	ASN AAT(GLY GGA(LYS A A A

	•				
GLN A A A 1440	THR C A 1500	71/82 1260 2112	TYR ' A C 1620	THR , C A 1680	GLX ; G T 1740
GLN C A 1	THR TAC 1	VAL 1560	F	THR (A C / 1680	GLY 5 G G 7 174(
GIN GIN CAACAA	ARG THR CGTACA 1500	GLU VAL GAAGTG 1560	GLY GGT	GLY THR GGAACA 1680	LEU
GLU MET ARG GAAATGCGCC 1430	LEU VAL GLU LEU GLY LYS ILE ARG LEU ASP TTAGTTAGGAAAATTCGCTTAGAT 1470 1480 1490	ASP G A T (ARG ASN THR GLY SER ILLE ASN PHE GLY ILLE CGTAACACGGGTAGTATCAACTTTGGTATT 1590 1600 1610	GIN ALA SER VAL LYS GIN ASP ASN PHE LEU CAAGCAAGTGTTAAACAAGATAATTTCTTG 1650 1670	TYR GLY THR SER VAL ASN LEU GLY TATGGTACGAGTGTCAATTTGGGT 1720 1730 1740
MET A T G (LEU TTA	SER ASN AGTAAT 1550	GLY ' G G T 1610	PHE TTC 1670	VAL 'G T C 1 1730
GLU GAAA	ARG C G C	SER A G T	TIT	ASN AAT	SER AGTG
GIN C A G G	ILE A T T	G G T	ASN A A C	ASP G A T	THR ACG
ARG C G T	LYS A A A A	ASN A A T	ILE ATC1	GLN C A A B	GLY G G T 1
LEU 1 TTAC 1420	GLY I G G A A 1480	PRO ILE ASN CCTATCAAT 1540	SER] A G T A 1600	LYS (A A A A A C 1660	TYR (TATG 1720
THRACT	LEU TTA	PRO C C T	G G T	VAL G T T	ASP G A T
SER A G C	GW GAG	ASP G A T	THR GLY ACGGG	Ser A G T	ASN A A T (
ASP G A T 1410	VAL G T T 1470	11.E A T T 1530	ASN A A C 1 1590	ALA G C A 1650	THR LYS ASN ASP ACGAAAAATGAT 1710
ALA G C T	LEUTTA	ARG ILE CGAATT 1530		GLN C A A	
SER	C GIN	A A A C	GLU GAA	TYR T A T	G G T
_ ₽	ASN SER	GLU : G A A 1520	LYS : A A A 1580	SER : A G T 1640	ALA G C T 1700
THR ACCG	ASN A A T	WAL G T C	LYS VAL LYS AAAGTCAA 1580	ILE SER ATTAG 1640	SER ILE ALA AGTATAGO 1700
ASN A A T A	TYR TAT	PHE GLU THR VAL GLU TCGAAACAGTCGA 1510 1520	LYS A A A	G G T	SER A G T
G G A .	TRP T G G 50	GLU GAA	TYR T A T 70	SER AGT 30	VAL G T A 30
ARG PHE GLU GLY ASN THR VAL CGCTTTGAAGGAAATACCGT 1390 1400	GLU GLY THR TRP TYR ASN SER GAAGGAACTTGGTATAATTC 1450 1460		ASP VAL VAL TYR LYS VAL LYS GLU GATGTCGTATATAAGTCAAAGAA 1570 1580	GLY THR GLU SER GLY ILE SER TYR GGTACAGAGTGGTATTAGTTAT 1630 1640	GLY ALA ALA VAL SER ILE ALA GGGGCGCAGTAAGTAGCT 1690 1700
PHE TTTG	GLY G G A	PHE TTC1	VAL GTC	THR A C A	ALA G C G
ARG C G C	GLU GAA	G G T	ASP G A T	GLY GGT	G G G (

1					
_ & 0	= 0	22/82	., 4 0	_ U 0	. 40
GLU G A A 1800	SER A G T 1860	HIS C A T 1920	SER T C A 1980	ASN A A C 2040	GLY G G A 2100
PHE TTT	GLY G G A	GLY G G T	GIN SER CAATCA 1980	TRP T G G	GLY G G T
VAL SER LEU GLY GLY ASN VAL PHE PHE GLU GTAAGTCTTGGTGGAAATGTTTTCTTTGAA 1780 1790	TYR T A C	ASN SER TYR TYR VAL GLY LEU GLY HIS A A C T C C T A T T A G G T C A T 1900 1920	TYR ASN ARG ASN LEU TYR ILE TATAACCGTAATTTATATATT 1960 1970	CLY G G T	PRO THR LYS GLY VAL LYS ALA SER LEU GLY GLY CCAACTAAAGGGGTTAAAGCAAGTCTTGGTGGA 2070 2070 2080 2100
VAL C T T 0	THR ; A C T 1850	VAL GLY ; T A G G A '	TYR 1 T A T 1970	PHE 7 T T T 2030	ALA SER CAAGT 2090
ASN A A T G	THR ACG?	VAL GTA	LEU TTA	255 1 0 1	ALA G C A
G G A D	ARG C G T	TYR TAT(ASN A A T	PHETT	LYS A A A (
LEU GLY TTGGT 1780	LYS A A G 0	SER TYR CCTAT	ARG C G T 0	ASP GATO	VAL GTT
LEU (CTTG	ASN TYR LYS ARG THR THR AACTATAAGCGTACGACT 1840 1850	SER 7 TCCT 1900	ASN 1 AACC 1960	ASP PHE ASP PHE GACTTTGATTT 2020	GLY GG G G G G G G G G G G G G G G G G G
SER AGTO	ASN A A C	ASN AAC1	TYR T A T	ASP G A C	LYS A A A
VAL G T A	SER	asn a a t	GLU GAA	ASN A A T	THRACT
ст. 3 с т 1770	SER T C C 1830	GLU G A A 1890	LEU GLU CTAGAA 1950	THR A C A 2010	PRO C C A 2070
LYS ASP AAAGAT(THR A C A	ASN GLU AATGAA 1890	PHE ALA TTTGCT	LYS THR ASN AAAACAAAT 2010	PHETTC
LYS A A A	ASP THR SER SER GATACATCCTCT 1830	VAL G T A	PHETTT	ILE A T T	TYR PHE TATTTC
РНЕ ТНК ТТАСТ 1760	~ ⊟	_ F	SER ASN GTAAC 1940	ຸ. ບ	ູ ບ
PHE TTT 1	LYS A A A	所 T.T.C	SER AGT	ASN GLY A A T G G (ARG GLY A G A G G (
TYR TAT1	ASN SER NACTCT)	GLY GGT	ILE A T T	G G T	ASN A A T
GLU PRO SAGCCCT 1750	ASN A A C	THR LEU GLY CTTTAGGT1 1870	ASN LYS ILE A A T A A A A T T J 1930	PHE LYS CLY TTAAAGGT 1990	SER LEU ASN GCCTTAATA 2050
GLU FG A G C 1750	ASP	THR I ACTT 1870	ASN 1 AATA 1930	PHE 1 TTTA 1990	SER 1 A G C C 2050
THR A C C (TYR TAC(VAL GTT?	THR TYR	LYS A A A T	ASN A A C P
TYR THR GLU PRO TYR PHE THE TATACCGAGCCCTATTTAC 1750 1760	ASN TYR ASP ASN SER LYS SEF A A C T A C G A T A A C T C T A A A G 1810 1820	ASN VAL THR LEU GLY PHE PRO AATGTTACTTTAGGTTTCCC 1870 1880	THR TYR ASN LYS ILE SER ASN ACCTATAATAAATTAGTAA 1930 1940	MET LYS PHE LYS GLY ASN GLY A T G A A A T T T A A A G G T A A T G G 1990 2000	TYR ASN SER LEU ASN ARG GLY TATAACAGCCTTAATAGAGG 2050 2060

			23/82			
GLY	G G T 2160	GLY TYR GGATAT 2220	CATC 2280	TYR T A T 2340	T A A T 2400	SER A G C 2460
VAL GLN	CAG	GLY GGA	វជ្ជិ ១ ១	GLU GAA	G G	LYS A A G
	TACAAACTAAGTGCAGATGTACAGGGT 2140 2150 2160	ALA G C A	GLY G G T	GLY SER ILE GLY PRO ASN ALA ILE TYR ALA GLU TYR GGTAGTATTGGACCTAACGCAATTTATGCCGAATAT 2310 2320 2330	VAL ILE GLY GTGATTGGT(2390	LEU ILE VAL PRO THR PRO PHE VAL SER ASP LYS SER TTAATTGTGCCAACTCCATTTGTGAGCGATAAGAGC 2430 2440 2450
ASP	GAT.	SER \ T C T (2210	THR ALA 1 C A G C G (2270	TYR 7 T A T (2330	VAL ILE 3 T G A T T 2390	SER 7 A G C (2450
ALA	G C A	ALA G C A	THR ACA	ILE ATT	VAL GTG	VAL GTG
SE	AGT	LYS A A A	TYR T A T	ALA G C A	ASP GATG	PHETT
na i	C T A	ALA G C A	THR ACTI	ASN A A C (SER TCT(PRO C C A T
LYS	A A A C 2140	SER P TCTG 2200	GIN THR CAAACT 2260	PRO A C C T A 2320	LYS LYS ILE SER SER ASP A A G A A G A T A A G T T C T G A T 2370 2380	THR 1 ACTC 2440
TYR	TAC	VAL G T A	TYR T A T	GCAC	ILE A T A	PRO C C A
TYR	T A C	VAL G T T	PHE TYR TTCTAT	ILE A T T	LYS A A G i	VAL G T G (
LYS	A A C A A A T A C 2130	TRP T G G 2190	PRO C C G 2250	SER A G T A 2310	LYS A A G 2370	11.E A T T 2430
ASIN	AAC	LEU CTC	LEU TTA	GLY G G T	PHE T T T	LEU T T A
ASP	G A T	HIS C A C	ARG C G T	TYR T A T	THR ACT	GUU GAG
SE	2120	ASP \ G A T 2180	LYS A A G 2240	ALA ' G C T 2300	GLY 1 G G T 2360	ALA : G C A 2420
PRO GLY	G G T	ARG AGAG	ASN LYS A A C A A (2240	GLY PHE ALA GTTTTGC 2300	GLY THR GLY 3 G T A C T G G '	ALA SER ALA CTAGCGC 2420
	CCA	ASP G A C A	GLY GGA	G G T	G G T	ALA G C T
ILE	A T T 10	LEU TTA(GLY PHE 3 G T T T T C 2230	ARG C G T G	SER AGTO	ILE ALA THR TCGCTACAC 2410
選	ACTA 2110	PRO I C C A T 2170	G G T T 2230	SER LEU P CATTAC 2290	G G T A 2350	ALA 1 G C T A 2410
VAL	CGAGTTACTATTCCAGGTTCT 2110 2120	PHE TYR PRO LEU ASP ARG ASP TTCTACCCATTAGACAGAT 2170 2180	ALA ASN GLY PHE GLY ASN LYS GCAAATGGTTTTGGAAACAA 2230 2240		GLY ASN GLY SER GLY THR GLY THR PHE GGTAATGGTAGTGGTACTTTT 2350 2360	
ARG	C G A	PHETTC	ALA G C A J	G G T	G G T	ALA G C A A

		24/82		
TRP T G G 2520	LYS A A A 2580	PRO C C A 2640	PHE T T C 2700	A A A 2760
LYS A A A	C C C	CLY G G G	GIN C A G	CTC
THR ACT/	TYR T A T	ILE A T T	GLU GAA	GAA
TRP ASN THR LYS TRP TGGAATACTAAATGG 2510 2520	ASP ' G A T' 2570	PRO 1 C C T 2630	VAL GTC (2690	1 T C A 2750
TRP T G G A 25	PRO C C T (SER TCT(ASP GAT	T C A
VAL G T T	U GLU SER ASP VAL LEU LYS ARG LEU PRO ASP TYR GLY LYS A G A G C G A T G T A T A A A A G A T T G C C T G A T T A G C A A A A G A T T G C C T G A T T A T G C C A A A 2580 2570 2580	GUN PHE GIN TRP GIN SER PRO ILLE GLY PRO GGATTCCAATGGCAATCTCCTATTGGGCCA 2640	ASP G A T	PHE *** *** TTCTAATAAATTGAACTTTTTTCTTCAGAACTCAAA 2730 2740 2750 2760
SER A G T (ARG A G A 1	TRP T G G (ASN A A T (T T T
ALA : G C A A 2500	LYS AAAAAA2560	GIN 7 C A A T 2620	GLU 7 GAAA 2680	CTTT 2740
ASP ALA ALA SER GATGCGGCAAGT 2500	LEU TTA	PHE TTC	TYR T A T	GAA
ASP G A T	VAL G T A	G G A	LYS A A A	ATT
VAL G T T 2490	ASP G A T (2550	VAL G T C (2610	LYS A A A 2670	*** T A A 2730
PHETIT	SER AGC	GLY GGT	ILE A T T	*** T A A
LEU TTA	GW GAG1	THR	PRO C C A A	
SER C T C C 2480	LEU 1 T T A 2540	ALA SER 3 C T C T 2600	TYR ALA LYS 'ATGCCAAA 2660	SER r r c r 2720
ARG THR SER GAACCTCO 2480	ASN GLY LEU A A T G G A T T	AIA G C C	ALA G C C	GGT!
ARG C G A	ASN A A T	ARG C G C G	TYR T A T	G G A G
VAL GTC 70	LYS AAAA 30	ILE ATT(90	SER TCT 50	ILE ATT(
GIN ASN THR VAL ARG THR SER CAAAATACGGTCCGAACCTCC 2470 2480	LYS SER ASP LYS ASN GLY LEA A A A T C A G A T A A A A A T G G A T T 2530 2540	SER SER ARG ILE ARG ALA SER TCAAGCCGTATTCGCGCCTCT 2590 2600	VAL PHE SER 3 T A T T C T C T 7 2650	SER] A G T A 2710
ASN A A T	SER T C A	SER AGC(VAL G T A	PHE T T T A
GIN C A A	LYS A A A 7	J C A	LEU VAL PHE SER TYR ALA LYS PRO ILE LYS LYS TYR GLU ASN ASP ASP VAL GLU GIN TTGGTATTCTCTTATGCCAAATTAAAAATATGAAATGATGATGATGAACAG 2650 2670 2680 2690	GIN PHE SER ILE GLY GLY SER CAATTTAGTATTGGAGGTTCT 2710 2720

ATTA	2820
CCCATCATTA	2810
AAATATAAA	2800
TGGGCAGAGA	2790
CCTAATTTAAT	2780
CAACGTTCTCTGC	2770
AAO	

TTGCAC	2880	-
TTGCTTTAGGTA	2870	
AACCGCACTT	2860	
TCGCAAAAGT	2850	
ATGAAAAACA	2840	
AGGATATTTATCAA	2830	

TATANTT	5/82 040 5
AATGCGGGT	2930
TGCTTTCATT	2920
AAGAAAAAT	2910
GCTTCCGCTG	2900
TGCTTCAGGCTAT	2890

TNCAAGGCNAAGG

SB33 D15

	IS THE SEQUENCE BEING TRANSLATED	
	GGCATTGAAAAAAGGACAGCTTTCCCTTTAACCTTGAAATATAGGGAAATT 10 20 30 40 50	A C T T 60
	ACTGGCGATTTGAATTAATTTAAGTGGGCCAATTTCTATTGCAAAAGGTGC	T G G C 120
	GCATCAGCAATTGGATTGGTGTATTTTTAAGTTTTATGGCATTAGTGT	86/82 180
	TTAGGGATTATTTTTTCCATTACCAGTATTAGATGGCGGTCATTTAGTTTT	ттта 240
•	ACAATGGAAGCTGTTAAAGGAAAACCTGTTTCTGAGCGGGTGCAAAGCATCTGTTAT	TCGA

		27/8	32		
С G Т 360	GLY G G 420	VAL G T 480	VAL G T 540	ASP G A 600	ILE A T 660
T. A	PHE TTC	GLX G G T	ARG C G T	ASP G A T	ILE A T C
∺ ÷	LEU T T A	ASP G A T	GIN C A G	PHE TTC	SER T C G
;	T. A	VAL GTG 10	GLY G G T	ARG C G A	LYS A A A
A T G A 350	SER 1 AGTT 410	ARG VAL CGTGTG 470	MLA (G C C G 530	GN 7 GGTC 590	ALA LYS G C T A A A 650
T T T	ALA G C A	ILE A T T	ARG C G T	SER AGT	VAL G T G
T W T	ILE A T C	ASP G A T	VAL GTT	VAL G T A	VAL G T T
340	LEU CTA	ALA LYS ASP GCAAAAGAT 460	PRO C C T 520	PHE T T C 580	SER A G C 640
5 -	LEU	ALA 3 C A	LEU	LEU TTA	VAL SER GTTAGC 640
- 9	LYS A A A		SER A G T	SER	VAL GTT
330	LYS A A A A 390	PHE VAL TTTGTG 450	ALA G C A 510	ARG C G C 570	LEU CTT 630
E	MET A T G	PRO C C T	ARG C G A	VAL GTC	VAL G T G
• •	T C G	ALA G C A	TLE A T C	ILE A T T	ASP G A T
c c	A	S C	NE A A	A T	E C
320	A T A C 380	РНЕ <i>1</i> ТТТG 440	GIN C CAAC 500	ALA 7 G C T A 560	GLU CGAAG
•	A G G	WAL GTG	GW GAA(ASP VAL	ALA HIS GIN CGCATCAAC
) :)	T A T	тн тн 1 с G A C T (430	LEU TTA(ASP GAT	HIS C A T
310	T T A 370	THR A C G 430	ASP G A C T 490	ASN A A T 550	ALA G C G 610
))	AAT	THR ACAP	GIN GLY	ASP G A C	LYS A A A G
310	СТАТААТТТАТАТА G G A T A C 370	THR THR THR THR VAL PHE I TACGACAACGACTGTGTTG 430	GIN GLY ASP LEU GLU GLN GLN T C A A G G T G A C T T A G A A C A C A	THR ASP ASN ASP VAL ALA I GACTGACAATGATGTGGCTA 550 560	VAL LYS ALA HIS GIN GIU (TGTGAAAGCGCATCAAGAAG 610 620
	ບ	H	E4	G	E

		28/8	12		
LEU T.T. 720	ALA G C 780	ILE A T 840	LYS A A 900	GLN C A 960	GLY . G G 1020
ASN A A C	PRE TTT	PRO C C T	ASP G A T	LEU TTA	GLU 3 A A (
GLN C A A	GLU PHE AL GAATTTGC 780	GW GAA	ASP G A T	THR	PHE I T T (
LYS A A A 0	ASN A A T	VAL GTTO	GLU 3 A A O	SER A G T	LYS A A A S
LEU LYS CTAAAA 710	LEU TTAA 770	THR ACCG	ASN GLU AATGAA 890	SER SER AGCAGT 950	ASN 1
	LYS A A A '	ALA G C A	ILE A T C /	SER AGT 1	GLY 3 G A 7
GU ALA GAAGCA	GLU GAAA	ASN A A C	GIN C A A D	VAL GTT	TRP r G G (
PRO C C T 700	ARG GLU LYS LEU ASN CGAGAAAATTAAAT 760	ARG TYR ASN ALA THR VAL CGCTATAACGCAACCGTT 820	LEU ILE GLN ILE TTAATTCAAATC 880	SER TCT(TRP LYS LEU TRP GLY ASN LYS PHE GLU GLY TGGAAATTATGGGGAAATAAATTTGAAGG 1000 1020
ILE ILE PRO PRO ATTATTCCACCT 690 700	ILE A T T	ARG C G C	LEU TTA	GLU SER GAATCT 940	LYS A A A ?
ILE A T T	LEU TTA	GLY GGT	ILE A T T	ASN A A C	TRP I G G /
ILE A T T 690	11.E A T T 750	N VAL GLY TGTAGGT 810	GU ILE GAAATT 870	LYS GLY ASN AAGGGGAAC 930	TRP T G G 7 990
TAACTCT	ASP G A T	SE A G	ALA G C T	LYS A A G	SER
ASN A A C	0 0 0	ALA G C A	ARG C G C	PHE TTC	ASP G A T
	LYS VAL AAGTT 740	HIS TYR SACTAT 800	~ €	_	_ ⊢
LYS A A A 6	LYS A A A 7	HIS CAC	ASN AATA	LEU THR TTAAC1 920	GIN 1 C A A C 980
ILE ATC?	PHETT	GW GAG	PRO C C A I	ALA SER ;CATCA1	LEU TTAC
VAL LYS STTAAAA 670	G G G T	VAL IYS GLU TAAAAGAGO 790	THR LEU CCCTA(850	ALA G C A	GLU GAA1
VAL G T T 670	ASN A A C G 730	VAL G T A 790	THR A C G 850	LEU T T G G 910	MET A T G G 970
ASP GATG	ALA G C T.	SER AGT(ASN A A T A	ALA LYS	GIN C A A A
SER ASP VAL LYS ILE LYS GL TTCAGATGTTAAAATCAAAGG 670 680	ASP ALA ASN GLY PHE LYS VAL AGATGCTAACGGGTTTAAAGT 730	GIN SER VAL LYS GLU HIS TYR CCAAAGTGTAAAAGAGCACTA 790 800	VAL ASN THR LEU PRO ASN ASN TGTCAATACGCTACCAAATAA 850 850	ALA LYS LEU ALA SER LEU THE A G C C A A A T T G G C A T C A T T A A C 910	GLU GIN MET GLU LEU GIN PRC AGAACAATGGAATTACAACC 970 980
H	X	υ	H	A	K

			29/82			
ALA	1080	VAL : G T 1140	ASN 1 A A 1200	THR 'A C 1260	ARG C G 1320	LEU . T. T 1380
TYR	TAT	ASN A A T G	GLY G G T	ASP G A T	GLU GAAC 13	THR ACA
GLY	ວອ	VAL 3 T T	ILE A T A	ASN A A T	GLY 3 G G	LYS A A A
ASIN GLLY	ATC	LYS A A A G O	ILE LTT	LEU ASN FTAAAT	LEU GLY TTGGG	ASN LYS THR LEU AATAAAACATT 1380
ASIN	A T A 1070	тнк I A C A A 1130	ARG ILE ILE CGCATTATA 1190	HIS LES CATTT 1250	LYS I AAC 1310	ALA P 3 C A A 1370
B	TAA	LYS A A A A	ALA G C A C	LEU TAC	ALA LYS 3 C A A A A C 1310	ASP 3 A T G
TY.	ATT	elu A A A	SER AGIG	ALA C A T	LYS A A A G	ASP 3 A C G
TYR TYR LEU	1060	ASP GLU ; A T G A A / 1120	ARG : G T A 1180	SER ALA LEU CAGCATTAC 1240	ILE \ T T A 1300	PHE . TTG 1360
ASP	CGTGATTATTTAATTAATGGCTATGC 1060 1070 1080	ASN A T G	LEU Z	LEU	ALA 3 C A A	PRO ASP PHE ASP ASP ALA ASN LYS THR LEX CCTGATTTTGACGATGCAAATAAAACATT 1360 1370 1380
ARC 1	G T G	LEU 1	ASP 1 ACC	LEU 1	ASN 1	PRO A
ILE ARG		GLN 1 CAGC 1110	GIN TYR ASP LEU ARG SER ALA ARG ILE ILE GLY ASP CAGTATGACCTTCGTAGTGCACGCATTATAGGTAA 1170 1180 1200	GLU LEU GLU PRO LEU LEU SER ALA LEU HIS LEU ASN ASP THE FAGCTTGAACCTTTACTTTCAGCATTACATTAAATGATAC 1230 1240 1250	GLU 71 G A A A A 1290	SER VAL PRO CTGTACCT 1350
ALA	C A A T 1050	VAL STIC	GIN (GLU 3 A A C 12	VAL 3 T A G	SER C T G
CIN	CAGGCAATT 1050	ASP 3 A T G	LEU (TTAC	LEU	ASP 3 A T G	AL ASN SER VAL TAAATTCTGTA 1350
		ALA A	GLY]	GLU 1 3 A G C	ALA S C A G	VAL 3 TAA
ASP	A T T 1040	LYS 7	CLU (3 A A G	ALA (3 C C G 1220	ILE A T T G 1280	THR CAG CAG 1340
LYS	AAG	CTA	ASN A T G	SER	ASP 3 A T A	THR CAA
CELU	AGA	ILE T C A	VAL TAB	MET TGT	SER AGTG	ASN A C A
PHE GLU	TCG 1030	GLN : A A A 1090	ASP VAL 3 A T G T A 1 1150	GLY 5 G T A 1210	ARG : G T A 1270	GLY: ; G T A 1330
GLN	AAT	ALA GIN ILE THR SCACAAATCACTA 1090	ILE ATTG	GLY GLY MET 3GAGGTATG1 1210	ARG G C C	GLY TYR GLY ASN THR THR GTTACGGTAACACAACA 1330 13
ALA	TGCGCAATTCGAGAAAGATTTG 1030 1040	LYS ALA GIN ILE THR LYS ALA ASP VAL GIN LEU ASN ASP GLU LYS THR LYS VAL ASN VAL CAAAGCACAAAACTAAAGCGGATGTTCAGCTAAATGATGAAAAAAAA	THR ILE ASP VAL ASN GLU A A C C A T G G T G T A A A T G A A G 1150	LEU GLY GLY MET SER ALA TCTGGGAGGTATGTCTGCCG 1210	PHE ARG ARG SER ASP ILE ALA ASP VAL GLU ASN ALA ILE LYS ALA LYS LEU GLY GLU AR TTTCCGCCGTAGTATTGCAGATGTAGAAATGCAATTAAAGCAAAACTTGGGGAACG 1270 1280 1280 1280	GLY TYR GLY ASN THR THR VAGGTTACGGTAACAG 1330
	£	ر ن ر	A	F.	E E	A C

		30/82			
GLU ' G A 1440	THR A A C 1500	PHE T T 1560	VAL : G T 1620	THR GLU 1 C A G A 1680	ALA ; G C 1740
PHE TTT 1	GLY G G A A 15	PE TTCT 15	VAL GTC	THR ACA	ALA G C G
NSP ALA GLY ARG ARG LEU THR VAL HIS GLN LEU ARG PHE GL A T G C T G G C G T T T A A C T G T T C A C C T T C G C T T T G A 1410 1420 1430	GIN GIN GLU SAACAAGAA(1490	ARG THR GLY :GTACAGGT 1550	ASP G A T	GLY TYR GLY GGTTACGGT, 1670	GLY THR GLY GAACAGGG 1730
LEU CTT(GLN C A A	THR ACA	GLU VAL 3 A A G T G 1610	TYR T A C 70	THR ACA 30
GIN L. C A A C 9	GLN G C A A C 1490	ARG 7 C G T A 1550	GAAG 1610	GY 1 GGTT 1670	GLY 1 G G A A 1730
HIS CACC	ARG C G C	ASP G A T	ASP GAT(ILE A T T	LEU TTGG
THR VAL CTGTT 1420	GLU MET 3 A A A T G 1480	ARG LEU ASP : G C T T A G A T (SER ASN GCAAT 1600	ИЕ GLY ILE ТТGGТATT(ASN PHE .ATTTC1 1720
тнк АСТ 1420	αгυ G A A 1480	ARG C G C 1540	SER A G C 1600	РИЕ ТТТ 1660	ASN A A T 1720
LEU TTAA	GIN C A G	ILE A T T	G G T	ASN A A C 1	ASP G A T
ARG CGTT	ARG CGTC	LYS A A A I	ASN A A T	SER ILE GTATC.	GIN C A A
ARG 1 C G A C 1410	LEU PTTA 1470	GLY A G G A 1530	ILE 1 A T C 1590	SER LAGT 1650	VAL LYS GIN 3 T C A A A C A A 0 1710
G G A	THR ACT	LEU	PRO C C T	G G T	WAL G T C
ALA G C T (SER A G T	VAL CLU STIGAG'	ASP G A T	THR A C G	SER AGT(
ASP GATG	ALA ASP CTGAT 1460	VAL G T T 20	ILE ATT 80	ASN AACA	N ALA . A G C A 1700
VAL VAL P GTTGTTG 1400	ALA 7 G C T G 1460	LEU V TTAG 1520	ARG C G A A 1580	ARG 1 C G T A 1640	GLN CAA 17
VAL G T T	SER TCT	ASN SER GIN A T T C A C A A 1 1510	ASN A A C	VAL LYS GLU STCAAAGAA 1630	ILE SER TYR TTAGTTAT 1690
THR PHE NCCTTT(1390	VAL GTT1	SER	VAL GLU 5 T T G A A 7 1570	LYS A A A	SER AGT
THR A C C 1390	THR A C C G 1450	ASN A A T 1510	VAL G T T 1570	VAL G T C 1630	ILE A T T 1690
ILE A T A A	ASN A A T	TRP TYR	THR ACA(LYS A A A (G G T P
ALA ILE THR PHE VAL VAL A AGCGATAACCTTTGTTG 1390 1400	GLY ASN THR VAL SER ALA ASP SER THR LEU ARG GLN GLU MET ARG GLN GLN GLU GLY THE AGGAAATACCGTTTCTGCTGATAGTACTTTACGTCAGGAAATGCGCCAACAAGAAGGAAC 1450 1460 1500	TRP TYR ASN SER GIN LEU VAL GIU LEU GLY LYS ILE ARG LEU ASP ARG THR GLY PHE PHE TTGGTATAATTCACAATTAGTTAGGAAAAATTCGCTTAGATCGTACAGGTTCTT 1510 1520 1530 1560	GLU THR VAL GLU ASN ARG ILE ASP PRO ILE ASN GLY SER ASN ASP GLU VAL ASP VAL VAL CGAAACAGTTGAAACCGAATTGATCCTATCAATGGTAGCAATGAAGTGGATGTCGT 1570 1580 1590 1600	TYR LYS VAL LYS GLU ARG ASN THR GLY SER ILE ASN PHE GLY ILE GLY TYR GLY THR GLA A T A T A A A A G A A C G T A A C A C G G G T A G T A T T G G T A T T G G T T A C G G T A C A G A A G 1640 1650 1650 1650 1660 1670 1680	SER GLY ILE SER TYR GLN ALA SER VAL LYS GLN ASP ASN PHE LEU GLY THR GLY ALA ALA GAGTGGTATTAGTTATCAAGCAAGTGTCAAACAAGATAATTTCTTGGGAACAGGGCGGC 1710 1720 1730 1740

ASS.

TRP ASN TYR

GLY

黑

SES SES

出

LYS THR ASN ASP PHE ASP

GLY ILE

ASIN

GEY

1990

2020

TAAAATTAGTAACTTTGCTCTAGAATATAACCGTAATTTATATATTCAATGAAATT

GIN SER MET LYS

TYR ILE

ASN LEU

ARG

S

TYR

ag B

园

出

NS.

SER

IE

FIG.1D.(CONTINUED)

		502	
GLU	ASP	THR r A C 1920	ASN
G A	G A		A A
1800	1860		1980
LYS ASN ASP TYR GLY THR SER VAL ASN LEU GLY TYR THR GLA A A A A A T G A T T A C G A G T G T C A A T T T G G G T T A T A C C G A 50 1770 1780 1800	GLY VAL SER LEU GLY GLY ASN VAL PHE PHE GLU ASN TYR ASP GGTGTAAGTCTTGGAAATGTTTTCTTTGAAACTACGA 20 1830 1840 1850 1860	VAL 3 T T 2	TYR LAT
TYR 'A T A	ASN A A C 1	ASN A A A A T G	ASN SER TYR VAL GLY LEU GLY HIS THR NACTCCTATTATGTAGGATTAGGCCATACC1 1950 1970
GTT	GLU A A A	SER GT	HIS
ы с	PHE G	GAA	GLY 1
ТСС	TTG	GAA	16 C C
1790	1850	1910	1970
SN I	治	IYR (LEU
	TCT	A T G	TAG
AL A TCA	AL ETTT	展 C T T	E A T
SER V. G.T. G.T. G.T. 1780	ASN VAL A A T G T T 1 1840	THR 1 C G A 1900	VAL (3 T A G 1960
SE A G 17	AS 1 18 1 18	TF	T G T
THR	G G 7	ARG	TYR
A C G		C G 1	T A '
GLY	GLY	LYS	TYR
G G T	G G T	A A G	T A T
TYR 1770	SER LEU GLY GLY AGTCTTGGTGGAP 1830	TYR : T A T 1890	SER : T C C 1950
ASP	SER	ASN	ASN
3 A T 7	AGT	A A C	AAC
LYS ASN ASP TYR GLY THR SER VAL ASN LEU GLY TYR THR A A A A A T G A T T A G G T A C G A G T G T C A A T T T G G G T T A T A C C 6	GLY VAL GGTGTA?	SER SER ASN TYR LYS ARG THR THR TYR GLY SER CCTCTAACTATAAGCGTACGACTTATGGAAGT1 1890 1900 1900 1910	GLU ASN SAAATA)
LYS A A A I	GLY 3 G T (SER T C C	GU GAA O
THR 1	ASP (THR S	
A C G A	3 A T G	ACAT	
1760	1820	1880	
GLY	LYS	ASP	VAL
3 G T A	A A A G	G A T	G T A
ALA	THR	SER	PRO
G C T G	ACT/	AGT	C C T G
ILE	TYR PHE THR	LYS	РНЕ
1 T A G	'ATTTTACT!	A A A A	ТТСС
1750	1810	1870	1930
VAL SER ILE ALA GLY THR	PRO TYR PHE THR LYS ASP	ASN SER LYS SER ASP THR SER SER ASN TYR LYS ARG THR THR TYR GLY SER ASN VAL THE	LEU GLY PHE PRO VAL ASN
AGTAAGTATAGCTGGTACG	GCCCTATTTTACTAAAGAT	TAACTCTAAAAGTGATACATCCTCTAACTATAAGGGACTTATGGAAGTAATGTAAC	TTTAGGTTTCCCTGTAAAT
1750 176	1810 18	1870 1870 1880 1890 1920	1930
VAL 3. T. A. A.	PRO	ASN	LEU
	CCT	A C 1	TA(
A G	ე ე	TAT	T.

31/82

SUBSTITUTE SHEET

		32/82			
THR 2 A C 2160	PRO 2220	. .	LEU 1 T T 2340	LYS ' A A 2400	GLU A G A 2460
VAL GTT?	TYR F A C	ASN AAT(SER TCAT 23	ASN A A T	ALA G C A
ARG C G A (PIE	ALA G C A	G G T	ASN A A T	SER AGCG
GLY GLY 5 G T G G A 2150	GIN GLY C A G G G T 7 2210	GLY TYR ALA 3 G A T A T G C A 2270	ILE A T. T 30	GIN C A A 30	ALA G C T
	TYR LYS LEU SER ALA ASP VAL GIN GLY PHE TACAAACTAAGTGCAGATGTACAGGGTTTC 2190 2200 2210	G G A T 2270	PHE TYR GIN THR TYR THR ALA GLY GLY ILE GLY SER LEI TTCTATCAAACTTATACAGCGGGTGGCATTGGTTCATT 2310 2320 2330	TYR GIN GLY GIN TATCAAGGTCAA!	ASN ALA ILE ALA THR ALA A A T G C A A T C G C T A C A G C T P 2440
LEU CTT(VAL G T A C	ALA G C A (G G T	GIN C A A	ALA G C T
ALA SER 5 C A A G T 0 2140	ASP G A T	SER	ALA G C G	TYR T A T	ILE A T C
ALA G C A 2140	SER ALA ASP G T G C A G A T C 2200	ALA G C A T 2260	THR A C A 2320	ILE A T T 2380	ALA G C A 2440
LYS A A A (Ser ag t	LYS A A A (THR TYR	ALA G C A J	ASN A A T
GLY VAL 3 G G G T T 130	C T A	ALA G C A	THR	ASN A A C G	CLY G G T
GLY \ G G G 2130	TYR LYS LEU 'ACAAACTA 2190	SER ATCTG 2250	TYR GIN FATCAA 2310	PRO 3 C C T 2370	SER ASP VAL ILE GLY GLY FCTGATGTGATTGGTGGT 1 2420
LYS A A A G 21	TYR TAC	VAL G T A '	TYR T A T	GLY GGGG 2.	ILE A T T
ACT	TYR A T A C	VAL G T T	PE	ILE A T T	VAL G T G
()	ASN LXS AACAAA 2180	LEU TRP T C T G G 2240	LEU PRO FTACCG 2300	SER AGC 60	ASP G A T 20
所E 1 エエCC 2120	ASN I A A C A 2180	LEU 1 CTCT 2240	LEU TTA 23	G G T A 2360	SER A TCTG 2420
TYR TAT1	ASP G A T	ASP HIS	ARG C G T	ALA TYR GLY SER CTTATGGTAGO 2360	ASN LYS ILE SER AATAAGATAAGT 2410
ARG GLY 1 G A G G C 1 2110	SER	ASP G A T	LYS A A G C	ALA G C T	ILE
A G A 2110	G G T T 2170	ARG A G A G 2230	ASN A A C A 2290	PHE TTT(2350	LYS A A G 2410
ASN A A T A	PRO C C A (ASP GACA	PHE GLY	G G T	ASN A A T
LEU ASN ARG GLY TYR FAE FR CCTTAATAGAGGCTATTTCCC 2110 2120	ILE PRO GLY SER ASP ASN LYS TYR A A T T C C A G G T T C T G A T A C A A T A C	LEU ASP ARG ASP HIS LÆU TRP VAL VAL SER ALA LYS ALA SER ALA GLY TYR ALA ASN GLY A T T A G A C A G A T C A C C T C T G G G T T G T A T C T G C A A A G C A T C T G C A A A T G G A A A T G G A A A T G G A A A T G G A T A T	PHE CLY ASN LYS ARG LEU PRO TTTTGGAAACAAGCGTTTACCG 2290 2300	ARG GLY PHE ALA TYR GLY SER ILE GLY PRO ASN ALA ILE TYR GLN GLY GLN ASN ASN LY. A C G C G T T T T G C T T A G C A T T G G G C C T A A C G C A A T T T A T C A A G G T C A A A T A A T A A A A A A A A A A A	PHE ASN LYS ILE SER SER ASP VAL ILE GLY GLY ASN ALA ILE ALA THR ALA SER ATTTAATAAGATAAGTTCTGATGTGATTGGTGGTAATGCAATCGCTACAGCTAGC 2430 2440 2450
9		~~	•		

		33/82			
LEU : C T 2520	GLU LG A 2580	THR ' A C 2640	PRO 1 C C 2700	PHE 7 T T 2760	T A 2820
SER TCC(LEU	SER	LYS A A A	SER T C T	ATT
THR A C C	G G A	ALA G C C	ALA G C T	G G T	CTA
ARG C.G.A.1	ASN A A T	ARG C G C C	TYR T A T (GLY G G G 60	T G C
VAL 1 G T C C 2510	LYS ASN A A A A A T 2570	THR PACTC 2630	SER 1 TCTT 2690	ILE (ATTG 2750	СТСТ 2810
THR ACA(ASP G A T	ARG C G T	PHE T T T	SER A G T	GT
ASN A A T	LYS TRP LYS SER ASP A A A T G G A A A T C A G A T 2560	1 PRO ASP TYR GLY LYS SER SER ARG THR ARG ALA SER THR ACCCGATTATGGCAAATCAAGCCGTACTCGCGCCTCTAC 2610 2620 2630 2640	1 SER PRO SER GLY PRO VAL VAL PHE SER TYR ALA A T C T C C T A G T G G T A T T T T C T T A T G C T 2670 2680 2690	PHE GIN PHE SER ILE GLY TTCCAATTTAGTATTGGG 2740 2750	GTCATCAGAACTCAAAACAACGTTCTCTGCCTAATTTA 2790 2820
GLN C A A A 2500	LYS A A A 2560	SER T C A 2620	VAL G T G 2680	GLN C A A 2740	A A C 2800
SER AGT	TRP T G G	LYS A A A	PRO C C A	ME T T C (AAA
LYS A A G	LYS A A A	TYR GLY FATGGC,	GLY G G A	GIN C A G	ပ် မ
ASP ' G A T 2490	ASN THR AATACT 2550	TYR T A T 2610	SER 7 A G T 2670	VAL GLU G T C G A A 0 2730	1 G A A 2790
SER AGT (ASN AAT	ASP G A T 1 26	PRO C C T	Ö	T C A
VAL GTG	TRP T G G	PRO C C C	SERTCT	ASP G A T	T C A
PHE TTT 0	_ E ·			<u>م</u> ا	_
PRO 1 C C A T 2480	SER VAL A G T G T 7 2540	ASP LEU G A C T T 2 2600	TRP GIN T G G C A 1 2660	ASN ASP A A T G A 1 2720	TTTT 2780
THR ACT(ALA G C A	LYS A A A G	GIN C A A	TYR GLU FATGAA	CTT
PRO C C A A	ASP ALA 3 A T G C G (2530	LEU TTG	PETCO	TYR T A T	G A A
VAL G T G C 2470	ASP G A T 2530	ASN VAL LEU ATGTCTTGP 2590	GLY G G A 7 2650	LYS A A A 2710	A T T 2770
ILE A T T (VAL G T T (ASN A A T	VAL GTCG	ILE LYS LYS TTAAAAAA 2710	*** T A A
LEU ILE VAL PRO THR PRO PHE VAL SER ASP LYS SER GLN ASN THR VAL ARG THR SER LEG GTTAATTGTGCCAACTCCATTTGTGAGTGATAAGAGTCAAAATACAGTCGAACCTCCT 2470 2480 2490 2520	PHE VAL ASP ALA ALA SER VA A T T T G T T G A T G C G C A A G T G T 2530 2540	SER ASN VAL LEU LYS ASP LE GAGCAATGTCTTGAAAGACTT 2590 2600	GLY VAL GLY PHE GIN TRP GLAAGGCA	ILE LYS LYS TYR GLU ASN AS A A T T A A A A A T A T G A A A A T G A A A Z T G A A A Z Z Z Z Z Z Z Z Z Z Z Z Z Z Z Z	*** *** CTAATAAATTGAACTTTTTC 2770 2780

山
\equiv
=
=
4
Ñ
\mathbf{O}
\sim
Τ.
(
Ī

	34/82
G C A G C T 2940	34102
C A G G C T A T G C T 2930	
3 A C T T G C T T 2920	
T T T A G G T A T T G (TAATGCAGGTTATA 2970
C C G C A C T T G C 2900	; Сттттаттаа 2960
ATCGCCAAAGTAA 2890	6 А А G А А А А А Т Т G С Т Т Т Т А Т 2950 2960
	TAACCGCACTTTAGGTATTGCACTTGCTTCAGGCTATGCTGCAG 2900 2930

FIG.1E

JB-1042-9-4 DNA, PAK D15 IS THE SEQUENCE BEING TRANSLATED

A A A A G G C A T T G A A A A A A A A A A A A A A A A	C A G G 20	гтссстттта 2 30	ACCTTGAAAA1 40	ACAACTTTCCCTTTAACCTTGAAATATTAGGGAAATT 30 40 50	
ACTTACTGGCGATTTC 70	G T C A T T A A A T A A T 80	PTTAAGTGGG	CAATTTCTA1	TAAATAATTAAGTGGCCAATTTCTATTGCAAAAGGTGC 100 110 120	
TGGTGCATCAGCAAATATTG 130	r a т т G G a т т G G т C	3 T A T T T T T A 1 150	160 160	GATTGGTGTATTTTAAGTTTTATGGCATTGATTAGTGT 150 170 180	35/82
АААТТТАСССАТТАТС 190	A A T T 200	1 T A C C A G T A '	F T A G A T G G C G C 220	TATTTCCATTAGTATTAGATGGCGGTCATTTAGTTTT 210 220 230	
TTTAACAATGGAAGCTGT	TAAA	LCCTGTTTCT	A G C G G G T G C A	GGAAAACCTGTTTCTGAGCGGGTGCAAAGCATCTGTTA	

		36/	82		
1 P A T 420	T G	I A G C 540	c G 600	I I G A 660	1 A A A 720
T.T.	ASE G A	C A	PHE A TTCG 600	SER TCG	C A
LEU TTA'	VAL G T G	ALA GLY 3 C T G G T	ARG C G A '	LYS A A A	LEU LYS
SER AGTT	ARG C G T	ALA G C T	GLY	ALA G C T	LEU
ALA G C A A 410	ILE A T T 470	ARG C G T G 530	SER A G T (590	VAL G T G (650	ALA G C A C 710
LEU ILE ALA SER LEU CTAATCGCAAGTTTA 410	ASP G A T. I	VAL GTT	PHE VAL	VAL GTT(THR GLU
LEU CTAA	LYS A A A	PRO VAL CCTGT1	PHE T T C	SER A	THRACT
LEU CTT(ALA G C A 460	LEU T T A 520	LEU PHE VAL SER GLY ARG PHE 1 TTATTCGTAAGTGGTCGATTCG 580 590 600	VAL G T T 640	PRO C C C A 700
LYS A A A	VAL 3 T G	LY VAL GIN GLY ASP LEU GLU GIN GIN 11.E ARG ALA SER LEU PRO VAL ARG ALA GTGTTCAAGGTGACTTAGAACAACAAATCCGAGCAAGTTTACCTGTGCT 490 530 530	SER I C I	ASP VAL LYS ALA HIS GLN GLU GLY ASP VAL LEU VAL VAL SER VAL VAL ALA LYS SER GATGTGAAAGCGCATCAAGAAGGCGATGTGCTTGTTGTTAGCGTTGTGCTAAATCGA 610 620 630 630	ILE \ T T
LYS A A A	PHE TTT	ALA G C A 7	ILE VAL ARG TTGTCCGC7 570	LEU CTT	VAL G T T A
MET G A T G 390	A PRO A C C T 450	ARG CGA 10	1 VAL TGTC 570	ASP VAL SATGTG 630	г ушк стст 690
T C G	ALA G C A	ILE ATC 53	ILE ATT 5	ASP GAT	ASN AAC
CAA	ALA G C C	GIN C A A	ASN A A T	C C C C C	G G T
ATA	PHE TTT(GIN C A A	ALA G C T	GW G	LYS A A A
A G G 380	VAL G T G 440	GLU GAA S00	VAL G T G 560	GLN C A A 620	ILE A T C 680
T A T	ACT	ASP LEU GLU BACTTAGA SO	ASN ASP VALA PARA PARA PARA PARA PARA PARA PARA P	AIA HIS GIN SCGCATCAA	VAL LYS ILE FTAAAAT (
TTA	THR ACG	ASP G A C	ASN A A T	ALA G C G	VAL G T T
A A T 370	THR A C A 430	GLY G G T 490	ASP G A C 7 550	LYS A A A 610	ASP G A T (670
TAT	THR A C G	VAL GIN GLY TTCAAGGT(THR ACTG	ASP VAL LYS FATGTGAAAG	ILE SER
ACGTCTATAATTTATATAGG 370	HEGLY THR THR THR TCGGTACGACAACGACTC 430	VAL G T T	RG VAL THR ASP ASN ASP VAL GTGTGACTGACAATGATGTG 550 550	ASP G A T	ILE A T T
A C	出してい	6 T 6	8G T O	SP A T C	LE T C A

37/82						
J P A T 780	ј Р АС 840			G G 20	, T ; C T 1080	L A TA 1140
ASN GLU NATGAA 18	VAL GLU P 3 T T G A A C 840	ASP G A T	SER THR LAGER OF 1960	PHE C TTTG 1020	ສ ວ ວ	VAL 1 G T T 114
ASN A A T	VAL G T T	GLU GAA	SER A G T	LYS A A A	ASN A A T	LYS A A A
LEU TTA/	ALA THR SCAACCO 830	ASN AAT	SER A G C	GLY ASN LYS GAAATAAA 1010	ASN A A T	LYS THR AAACA 1130
LYS A A A T 770	ALA G C A 830	ILE A T C 890	SER A G T 950	GLY G G A 1010	LEU T T A 1070	LYS A A A 1130
GLU GAAA	ASN A A C C	GIN C A A	VAL G T T	TRP T G G	TYR T A T	GLU GAAA
ARG C G A G	ARG TYR CGCTAT	ILE A T T	SER T C T	LEU TTA	TYR T A T	ASN ASP AATGAT 120
LEU ILE TAATTO 760	ARG C G C 820	LEU T T A 880	GLU G A A 940	TRP LYS LEU TRP FGGAAATTATGC 1000	ASP G A T .060	ASN A A A T 1120
LYS VAL GLY ASP VAL LEU ILE ARG GLU LYS LEU ASN GLU E A A A G T T G G C G A T G T T T T A A T T C G A G A A A A T T A A A T G A A A A T T A A A T G A A A T T A A A T G A A A T T A A A T G A A A A	GLY G G T	ASN ASN ARG ALA GLU ILE LEU ILE GIN ILE ASN GLU ASP A A A T A A T C C C C C C C C A A T T C A A A T C A A T G A A G A T G 870 880 890	LEU THR PHE LYS GLY ASN GLU SER VAL SER SER THR L TTAACTTTCAAGGGAACGAATCTGTTAGTAGCAGTACAT 930 940 950	GIN PRO ASP SER TRP TRP LYS LEU TRP GLY ASN LYS PHE C CAACCTGATTCTTGGTGGAAATTATGGGGAAATATG 990 1000 1000	LEU GIN ALA ILE ARG ASP TYR TYR LEU ASN ASN CTGCAGGCAATTCGTGATTATTTTAAATAAT 1050 1060 1060	LYS THR ASP VAL GLN LEU ASN ASP GLU LYS THR LYS VAL AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
GLY ASP VAL GCGATGTTT 750	SER VAL	GLU GAA	G G G	SER TRP ICTTGGT)	ILE A T T	GLN C A G
ASP GAT	SER AGT	ALA G C T	C A A G 930	» SER ТТСТ 990	ALA G C A 50	VAL GTT
сти н СССС 750	ALA S GCAA 810	ARG CGT 87	路 TTC 93	ASP G A T T 990	GLN CAG	ASP V G A T G 1110
VAL G T T	TYR TAT(ASN A A T	THR	PRO C C T (LEU CTG	THR ACG
LYS A A A (ASN A A T	LEU TTA	GIN C A A	ASP G A T C	
PHE T T T 740	GLU GAG 800	PRO C C A 860	ALA SER SCATCA 920	GLU LEU SAATTA 980	LYS A A A 1040	THR A C T 1100
G G G	LYS GLU AAAGA(80	LEU CTGC	ALA G C A	GLU GAA	GAGAA 1040	ILE A T C
ASN A A C	VAL G T A	ASN THR AATACGO 350	LEU TTG	MET A T G	PHE T T C	GIN C A A
SN LEU ASP ALA ASN GLY PHE ACTTAGATGCTAACGGGTTT 730 740	HE ALA LYS SER VAL LYS GLU T T G C C A A A G T G T A A A A G A G 790 800	RO ILE VAL ASN THR LEU PRO CTATTGTCAATACGCTGCCA 850 850	SP LYS ALA LYS LEU ALA SER A T A A A G C A A A A T T G G C A T C A 910 920	EU GIN GIU GIN MET GIU LEU TACAAGAACAAATGGAATTA 970 980	LU GLY ALA GLN PHE GLU LYS A A G G T G C G C A A T T C G A G A A A 1030 1040	YR ALA LYS ALA GLN ILE THR A T G C C A A G C A C A A T C A C T 1090 1100
ASP G A T	LYS A A A	WAL GTCA	ALA G C A	GLU GAA	ALA G C G	LYS A A A
SN LEU ACTTA(ALA G C C A	RO ILE CTATT(LYS A A A	EU GIN TACAA(IU CLY AAGGTO	ALA G C C
SN C	T T	RO C T	SP A T	EU T A	LU A A	YR A T

		38/8	32		
E G : A G 1200	an A A T G 1260	JY G S G G 1320	S T A A 1380	ARG P G C T 1440	A A G
ILE A T 7	ASN A A T	GLY G GGGG 1320	LYS A A P 13	ARG C G (GLU GA 1
ILE A T T	LEU T T A	LEU CTT	ASN A A T	LEU CTT	GLN C A A
ALA ARG ILE ILE G CACGCATTATAG 1190	LEU SER ALA LEU HIS LEU TTTCAGCATTACATTA1 140	ALA LYS LEU 3 C A A A A C T T (ASP ALA ASN FATGCAAAT 1370	ARG GIN LEU: GCCAACTT 1430	ARG GLN GIN GLU G GACAACAAGAAG 1490 1500
ALA G C A (1190	LEU T T A 1250	ALA G C A 1310	ASP G A T 1370	ARG C G C 1430	ARG C G A 1490
SER A G T	ALA G C A	LYS A A A A	ASP G A C (VAL G T T	MET A T G
ARG C G T	SER T C A	ILE A T T	PRE TTT	THRACTO	GUU GAAA
LEU C T T (U LEU ACTT	ALA T G C A 1300	ASP F G A T ' 1360	LEU THR VAL FTTAACTGTT 1420	GIN C A G (
LEU GIN TYR ASP LEU ARG SER TACAGTATGACCTTCGTAGTO 1170 1180	LEU TTA	ASP VAL GLU ASN ALA ILE LYS SATGTAGAAATGCAATTAAA(1290	SER VAL PRO ASP PHE ASP CTGTACCTGATTTTGAC(ARG C G T	SER THR LEU ARG GLN GLU MET GTACTTTACGTCAGGAAATGO 1470
TYR T A T	GLU PRO LEU SAACCTTT?	GLU GAA	VAL G T A	ALA GLY ARG SCTGGACGA(1410	LEU TTA
GIN C A G	GU 3 A A 3	VAL 3 T A	SER CCT	GLY 3 G A	THA A C T
LEU G TTAC 1170	LEU CTTG 1230	ASP V G A T G 1290	ASN S A A T T 1350	ALA G C T (141(SER 1 AGTA 1470
SN VAL THR ILE ASP VAL ASN GLU GLY LEU GLN TYR ASP LEU ARG SER ALA ARG ILE ILE (A T G T A A C C A T G A A A T G A A G G T T T A C A G T A T C G T A G T G C A C G C A T T A T A G T G T A G T G A T A T A T	LY ASN LEU GLY GLY MET SER ALA GLU LEU GLU PRO LEU LEU SER ALA LEU HIS LEU ASN A GTAATCTGGGAGGTATGTCTGCCGAGCTTGAACCTTTACTTTCAGCATTACATTAAATG 1210 1220 1230 1230	SP THR PHE ARG ARG SER ASP II.E ALA ASP VAL GLU ASN ALA II.E LYS ALA LYS LEU GLY G A T A C T T T C C G C C G T A G T A T G T A G A A A T G C A A T T A A A G C A A A C T T G G G G 1270 1270 1280 1280 1280	LU ARG GLY TYR GLY ASN THR THR VAL ASN SER VAL PRO ASP PHE ASP ASP ALA ASN LYS A A C G A G G T T A C A C A C A G T A A T T C T G T A C T G A T T T G A C G A T G C A A A T A A A A 1380 1330 1370 1380	HR LEU ALA ILE THR PHE VAL VAL ASP ALA GLY ARG ARG LEU THR VAL ARG GLN LEU ARG C CATTAGCGATAACCTTTGTTGATGCTGGACGACGTTTAACTGTTCGCCAACTTCGCT 1390 1400 1410 1410	HE GLU GLY ASN THR VAL SER ALA ASP SER THR LEU ARG GLN GLU MET ARG GLN GLU G TTGNAGGAAATACCGTTTCTGCTGATAGTACTTTACGTCAGGAAATGCGACAACAAGAAG 1450 1450 1460 1460 1470
GW GAA	ALA G C C	ILE A T T	THR ACA	VAL GTT	ALA G C T
ASN A A T 1160	SER T C T 1220	ASP G A T 1280	THR A C A 1340	VAL G T T 1400	SER T C T 1460
VAL 3 T A	MET A T G	SER A G T	ASN A A C	PE LTT	VAL 3 T T
SN VAL THR ILE ASP VAL A T G T A A C C A T T G A T G T A 1 1150	LY ASN LEU GLY GLY MET GTAATCTGGGAGGTATC 1210	ARG ARG SER SGCCGTAGTO 270	LU ARG CELY TYR GELY ASN THR A A C G A G G T T A C G G T A A C A C A 1330 1340	HR LEU ALA ILE THR PHE CATTAGCGATAACCTTTG 1390	ASN THR VAL VATACCGTT ISO
R 11.E CATT 1150	GLY G G A 1210	ARG C C C C 1270	TYR 7 T A C 1330	ILE A T A 1390	ASN A A T 1450
THR ACC 1	LEU CTG	SP THR PHE ATACTTICO	GGT 1	ALA G C G	HE GLU GLY TTGAAGGA?
VAL 3 T A	ASN A A T	THR A C T	ARG 3 G A	LEU PTA	GLU 3 A A
A T (6 T 7	SP AT1	LU A A (HR C A 1	胎 7.7.(
1					

			′8 <i>2</i>		
л Р	ASP V	.Y Т	GLY A	YR T	ASN T
3 Т Т	3 A T G	3 Т А	3 G G G	A T A	A A C T
1560	1620	1680	1740	1800	1860
GLY	ASP	G G	G G G	TYR	ASN
G G	G A 1	1		T A 1	A A
ARG THR GLY : G T A C A G G T 7	GLU VAL	GLY TYR CLY T GTTACGGTA 1680	GLY THR	LEU GLY TYR T PTGGGTTATA 1800	PHE GLU
ARG	GW	GLY	GEY	LEU	PAE
	3 A A	3 G T	3 G A	r T G	TTT
ASP	ASP	11.E	LEU	ASN	PHE
3 A T (3 A T G	A T T G	F T G C	A A T 7	I T C 1
1550	1610	1670	1730	1790	1850
ARG LEU	ASN A A T (PHE CLY	PHE T C 7	SER VAL	ILE 4 T T '
ARG	SER ASN	PHE	ASN	SER	ASN
	AGCAAT	LTT(A A T	A G T (A A T
LY THR TRP TYR ASN SER GIN LEU VAL GLU LEU GLY LYS ILE ARG LEU ASP ARG THR GLY 1 GAACTTGGTATAATTCACAATTAGTTGAGTTAGGAAAATTCGCTTAGATCGTACAGGTT 1510 1520 1530 1560	ARG II.E ASP PRO II.E ASN GLY SER ASN ASP GLU VAL ASP V CGAATTGATCCTATCAATGGTAGCAATGATGAAGTGGATG	ARG ASN THR GLY SER ILE ASN PHE GLY ILE GLY TYR GLY 1 CGTAACACGGGTAGTATCAACTTTGGTATTGGTTACGGTA 1650 1660 1660	GLN THR SER ILE LYS GLN ASP ASN PHE LEU GLY THR GLY 1 CAAACAAGTATTAAACAAGATAATTTCTTGGGAACAGGGG 1710 1720 1730	LA ALA VAL SER ILE ALA GLY THR LYS ASN ASP TYR CLY THR SER VAL ASN LEU GLY TYR ' CGGCAGTAAGTATAGCTGGTACGAAAATGATTATGGTACGAGTGTCAATTTGGGTTATA 1750 1750 1760 1770 1770	ASP GLY VAL SER LEU GLY GLY ASN ILE PHE PHE GLU ASN 1 GATGGTGTAAGTCTTGGTGGAAATATTTTCTTTGAAAACT 1830 1840 1850
LYS A A A A	ASN A A T (ILE ATC	GIN C A A (GGT 1	G G T (
GLY	ILE	SER	LYS	TYR	LEU
G G A	A T C	A G T	A A A		CTT
LEU TTAC	РВО С С Т 10	GLY G G T	ILE ATT .0	ASP GAT	SER AGTC
VAL GIU I	ASP	ASN THR CLY	SER]	ASN A A A A T G	VAL S
TTGAGT	G A T C	VACACGGGT1	AGTA		G T A A
1530	1590	1650	1710		1830
VAL	ILE	ASN	THR	LYS	G G T
G T T	A T T	A A C	ACA	A A A	
LEU	_	ARG	GLN	THR	ASP
TTA(C G T 1	C A A A	ACG1	GAT(
GIN C A A 1520	HE PHE GLU THR VAL GLU ASN TCTTCGAAACAGTTGAAAAC 1570 1580	AL VAL TYR LYS VAL LYS GLU TCGTATATAAAGTCAAAGAA 1630 1640	HR GLU SER GLY ILE SER TYR CAGAGAGTGGTATCAGTTATO 1690 1700	G G T 1760	HR GLU PRO TYR PHE THR LYS CCGAACCCTATTTTACTAAA 1810 1820
SER	VAL GLU	LYS	SER	ALA	THR
TCA		A A A	A G T	G C T	ACT
ASN	VAL	VAL	ILE	ILE	PHETTT
A A T	G T T	G T C	A T C	A T A	
TYR	PHE GLU THR	LYS	GLY	SER	TYR
G T A T	TCGAAACAG	F A A A	C G T	1 A G T	C T A T
1510	1570	1630	1690	1750	1810
TRP	GLU	TYR	SER	VAL	024
T G G	GAA	TAT	A G T	G T A	C C C
LY THR TRP TYR ASN SER GIN	PHETTC	AL VAL TYR LYS VAL LYS GLU	HR GLU SER GLY ILE SER TYR	LA ALA VAL SER ILE ALA	HR GLU PRO TYR PHE THR LYS
GAACTTGGTATAATTCACA?		TCGTATATAAAGTCAAAGA!	CAGAGAGTGGTATCAGTTA1	CGGCAGTAAGTATAGCT	CCGAACCCTATTTTACTAA
1510 1520		1630 1640	1690 1700	1750	1810 1820
LY G A	五 T C T	AL T C	C A	LA C G	C C

E E

G Z

3

S

ALA

GLY VAL LYS

THR LYS

82

出

TYR

Z E

ARG BC

ASIN

SER LEU

8

		4	0/82
20 c	T T 80	er L FGA 2040	4 4 0
ASN (A A T G 1920	THR JACCT 1980	MET ATG 204	, TAT 210
SER	HIS	SER	ASN
A G T	C A T	T C A	A A C
G G A	CLY	GIN	TRP
	G G C	C A A	F G G
CTTATG 1910	LEU T T A (ILE ATT 2030	GLY 3 G T 7 2090
THR	GLY	TYR	PHE
	3 G A C	FAT1	[TT(
THR	VAL	LEU	SER
C G	3 T A (ICTI
THR SER SER ASN TYR LYS ARG THR THR TYR GLY SER ASN V	ASN GLU ASN ASN SER TYR TYR VAL GLY LEU GLY HIS THR T	ALA LEU GLU TYR ASN ARG ASN LEU TYR ILE GIN SER MET L	LYS THR ASN ASP PHE ASP PHE SER PHE GLY TRP ASN TYR A COURAD AS A CTTTGATTTTCTTTTGGTTGGAACTATA N 2000 2000 2100
A C A T C C T C T A T A A G C G T A C G A C T T A T G G A A G T A T G	A A T G A A A A A A C T C C T A T T A T G T A G G A T T A G G C C A T A C C T	GCTCTAGAATATACCGTAATTTATATATTCAATGA	
1890 1920	1950 1960 1960 1960	2010 2020 2030	
LYS A A G (TYR [A T 7	ARG G T 7	ASP 1 A T 1 20
TYR	SER	ASN	PHE
FAT1	1 C C J	A A C C	TTC
ASN	ASN	TYR	ASP
A A C 1	A A C 1	rata	3 A C T
SER A	ASN	GLU T.	ASN
ICTA	A A T 7	3 A A T 7	1 A T G
1890	1950	2010	2070
SER	GW	LEU	THR
	GAA	CTA(ACA1
THR	ASN	ALA	LYS
ACA	A A T	3 C T (A A A A
	VAL G T A 1940	PHE T T T 2000	ILE A T T 2 2060
SER A G T	PRO C C T	ASN A A C	ASN GLY ILE AATGGCAT1 2060
LYS	PHE	SER	ASN
A A A	TTC	AGT	A A T
ASN SER ACTCT 1870	GLY G G T 930	ILE A A T T 1990	G G T 2050
YR ASP ASN SER LYS SER ASP	AL THR LEU GLY PHE PRO VAL ASN GLU ASN ASN SER TYR TYR VAL GLY LEU GLY HIS THR 1	YR ASN LYS ILE SER ASN PHE	YS PHE LYS GLY ASN GLY ILE
ACGATAACTCTAAAAGTGAT	TTACTTTAGGTTTCCCTGTAAATGAAATAACTCCTATTATGTAGGATTAGGCCATACCT	ATAATAAATTAGTAACTTT	A A T T T A A A G G T A A T G G C A T T
1870 1880	1930 1940 1950 1960 1960 1960	1990 2000	2050 2060
ASP	THR	ASN	PHE
G A T	ACT	A A T	TTT
YR	AL	YR	YS
A C	T T	A T	A A

> ₀	T T 20	
ARG C G A G 2160	PHE ' T T C T 2220	
G G A	GLY G G T	-
G G T	GIN C A G	
LEU CTT 2150	VAL 3 T A (2210	
SEK G T (ASP 3 A T (
ALA 3 C A 1	ALA 3 C A O	
ATTTCCCAACTAAAGGGGTTAAAGCAAGTCTTGGTGGACGAG 2120 2150 2150 2160 2160 2160	ASP ASN LYS TYR TYR LYS LEU SER ALA ASP VAL GLN GLY PHE T 3 A T A A C A A A T A C A A A C T A A G T G C A G A T G T A C A G G G T T T C T 2180 2210 2220	
TT 7 21	LEU CTAA	
9 9 9 e	LYS A A A C	
AAG	TYR 'ACA	Cox Cox
C T A 2130	TYR TYR FACTAC 2190	į
CAA	LYS AAT	£
J C C	ASP ASN LYS 3 A T A A C A A A 2180	2
A T T	ASP 7 A T A 2180	Ç
C)	AL THR ILE PRO GLY SER A TTACTATTCCAGGTTCTG 2170	ACD UT
ACAGCCTTAATAGAGG	GLY S	r car
TA	0 V	
T A A 2110	PRO T C C A 2170	ACD
C T	ILE A T 1	[H.
A G C	THR ACT	DRO
A C	AL, T.T.	g

		41/8	2		
SN GLY PHE GLY ASN LYS ARG LEU PRO PHE TYR GLN THR TYR THR ALA GLY GLY ILE GLY S A T G G T T T T G G A A C C A T C C G T T C T A T C A G C G G G G G G G T G G C A T T G G T T T G T T G T T G T T G T T G T T G T T G T T G T T G T T G T T G T T G T T G T T T G T T T G T T T G T T T G T T T G T T T G T T T G T T T G T T T G T T T G T T T G T T T G T T T G T T T G T T T G T T T G T T T T G T T T T G T T T T G T T T T G T T T T G T T T T G T T T T G T T T T G T T T T G T T T T T G T T T T T G T T T T T T G T T T T T G T	GLY A G G T A 2400	SN GLY THR PHE ASN LYS ILE SER SER ASP VAL ILE GLY GLY ASN ALA ILE THR THR ALA S A T G G T A C T T T T A A T A A G T T C T G A T G T G T G G T A A T G C A A T C A C A A C T G C G A 2410 2420 2460	ER ALA GLU LEU ILE VAL PRO THR PRO PHE VAL SER ASP LYS SER GLA ASN THR VAL ARG T GTGCAGAACTTATTGTACCAACTCCATTTGTGAGTGATAAAAGCCAAAATACAGTCCGAA 2470 2520	HR SER LEU PHE VAL ASP ALA ALA SER VAL TRP ASN THR LYS TRP LYS SER ASP LYS ASN G CCTCCCTATTTGTTGATGCGGCAAGTGTTTGGAATACTAAATGGAAATCAGATAAAATG 2530 2530 2530 2540 2550 2550	ARG A C G C G 2640
ILE A T T (HIS C A T	THR	THR VALACE OF CARACT CA	LYS A A A	ILE A T T (
G G C 1	ALA GLU 5 C C G A A 2390	ILE THR LTCACAA 2450	THR A C A	ASP G A T	ARG CGTA
GLY G G T G 2330	л А ІА ТССС 2390	II.E A T C 2450	ASN A A T 1 2510	SER T C A G 2570	SER A G C C 2630
ALA G C G (TYR T A T	ALA G C A I	GIN C A A	LYS A A A	SER T C A
THR ACAG	ALA ILE 3 C A A T T 1 380	ASN A A T	SER AGC	TRP T G G	LYS A A A
TYR T T A T A 2320	PRO ASN ALA ILE TYR ALA GLU HIS CCTAATGCAATTTATGCCGAACAT 2380 2380	VAL ILE GLY CLY ASN STGATTGGTGGTAATC	LYS L A A A A 2500	ASN THR LYS TRP LYS AATACTAAATGGAAA 2560	AS ASP LEU PRO ASP TYR GLY LYS SER SER ARG ILLE A A G A C T T A C C T G A T T A T G G C A A T C A A G C C G T A T T 2610 2620
THR ACTI 23	ASN A A T G 23	G G T	ASP GATA 25	ACT	TYR , T A T
GIN C A A I	ILE GLY PRO TTGGGCCT 2370	ILE	SER	ASN 3 A A T	ASP
Е ТYR ГСТАТ 2310	SER ILE GLY AGTATTGGG 2370	VAL GTG	E VAL FTGTG 2490	M TRP TTGGA 2550	ASP LEU PRO 3 A C T T A C C T C 2610
所 : T T C 23	ILE ATT 23	ASP GATG 2430	PHE 1 T T G 2490	VAL r g T T 1 255(LEU CTT2
PRO C C G	SER	SER	PRO	SER A A G T	ASP AGA(
LEU TTA(GLY GGT	SER AAGT7	THR AACT)	ALA 3 G C A ,	LYS 3 A A A C
ARG 3 C G T 2300	TYR F T A T 2360	1LE 3 A T P 2420	PRO A C C A 2480	ALA I G C G 2540	LEU CTTG 2600
LYS CAAGO	ALA FGC 1	LYS T A A (VAL TGT	ASP TGA'	C G T
ASN A A A C A	PRE	ASN TAA	ILE TAT'	VAL TGT	C A A
GLY T G G A A 22290	ARG GLY PHE ALA TYR CGGGTTTTGCTTAT 2350 2360	PHE T T T 2	LEU ACT1 2470	A T T 2530	J SER G A G 2590
PHE TTTTG 22		SN GLY THR PHE ASN LYS ILE A T G G T A C T T T T A A T A G A T 1 2420	ALA GLU LEU ILE VAL CAGAACTTATTGTAC 2470	SER LEU PHE VAL ASP ALA CCCTATTTGTTGATGCG 2530	LY LEU GLU SER LYS VAL LEU I GATTAGAGCAAGGTCTTGA 2590
SN GLY ATGGT1	ER LEU CATTA(GLY T G G	ER ALA GTGC	HR SER CCTC	ATT
S A	E O	Z 4	ရ မ	光 い	či S

GIN TRP GLN SER PRO ILE GLY PRO LEU VAL PHE SER TYR A CAATGGCAATCTCCTATTGGACCATTGGTATTTCTTATG 2670 2680 2690 2700	GLU ASN ASP VAL GLU GIN PHE GIN PHE SER ILE GLY G GAAAATGATGTCGAACAGTTCCAATTTAGTATTGGG 2730 2740 2750 2760
PRO 1	PHE (
T G G A 2680	GIN A C A G 7 2740
TATT	VAL GLU STCGAP
PRO T C C T	VAL T G T
SER A T C 1 670	P ASP T G A T 2730
GLN S C A A	ASP G A T
TRP	ASN A A T
GIN C A A	GLU GAA
PHE T T C 2660	
CLY GGA	LYS A A A A
WAL G T C	LYS A A A
GLY VAL A G G T G T C 2650	ILE \ A T T 2710
TACA	PRO C C A
LA SER, THR GLY VAL GLY PHE CCTCTACAGGTGTCGGATT 2650 2661	LA LYS PRO ILE LYS LYS TYR CTAAACCAATTAAAAAATAT 2710
C C	C T

42/8	32
. T T C T C T G C C	ттатсааат
C A A A A A C G A C 6	A T T A A G G A T A 1 2870
CTTTTTCGTCATCAGAACTCAAAAACGACGTTCTCTGCC 2790 2820	AATATTAAACCCATCATTTAATTAAGGATATTAATCAAAT 2850 2850 2860 2870
ACTTTTTCGT 80 2790	
LY SER PHE *** *** G C T C T T C T A A T A A A T T G A A 2770 2780	TAATTGAATTGGGCAGAGAA 2830
R PHE *** :TTTCTAA 2770	TGAATTG 2830
LY SER GCTC'	TAAT

TATGC	294
CTTGCTTCAGGC	1
AGGTTTTGCAC	2920
CTTGCTTT	2910
AGTAACCGCA	2900
GAAAACATCGCAAA	2890

TTATATTTTCAA	2980
CATTAATGCAGG	2970
AAAATTGCTTT	2960
T T C C G C T G A A G A A A	2950
-	

FIG. 1F

1. cad15 3. minnad15 2. eagand15 4. pakd15 5. sb33d15 1 minnad15 1 eagand15 1 pakd15 1 sb33d15 1	(1-2949) 5 (1-2953) 5 (1-2984) (1-2989)				ACAGGACAGCTTTCCCTTTTAACCTTGAAATATTAGGGA	1 aaaaGGCATTGAAAAACAGGACAACTTTCCCTTTTAACCTTGAAATATTAGGGA	Agctricccritinaccricaaaa	aaaaggcattgaaaaaaaaggacagctttcccttttaaccttgaaaatattaggga	
1. cad 3. min 2. eag 4. pak 5. sb3 minnad15 eagand15 pakd15 sb33d15 sb33d15	15 nad15 and15	3d15	₽	Н	~		. •		
		4. par 5. sb3	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	

43/82

cad15

6 CTTaCTGGCGATTTGTCATTAAATAATTTAAGTGGGCCAATTTCTATTGCAAAAGGTGCTG		2 CTTACTGGCGATTTGTCATTAAATTTTAAGTGGGCCAATTTCTATTGCAAAAGGTGCTG		cttactggcgatttgtcattaaataatttaagtgggccaatttctattgcaaaaggtgctg	44/82	7 GCaCATCAGCAAATATTGGATTGGTGTATTTTTAAGTTTTTATGGCACTGATTAGTGTAAA		3 GtGCATCAGCAAATATTGGATTGGTGTATTTTTTAAGTTTTTATGGCATTGATTAGTGTAAA		gcgcatcagcaaatattggattggtgtattttttaagttttatggca-tgattagtgtaaa
9	46	62	28		₩.	67	107	123	119	
minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

gattac 	TTTAGGGATTATGAATTTATTTCCATTACCAGTATTAGATGGCGGT	TTTAGGGATTATGAATTTATTTCATTACCAGTATTAGATGGCGGTCATTAGTTTTTA		TTTAGGGATTATGAATTTATTTCCATTACCAGTATTAGATGGCGGTCATTTAGTTTTTA	tttagggattatgaatttatttccATTACcagtattagatggcggtcatttagttttttta	15/82	ACAATGGAAGCTGTTAAAGGAAAACCTGTTTCTGAGCGGGTGCAAAGCATCTGTTATCGAA	ACAATGGAAGCTGTTAAAGGAAAACCTGTTTCTGAGCGGGTGCAAAGCATCTGTTATCGAA		ACAATGGAAGCTGTTAAAGGAAAACCTGTTTCTGAGCGGGTGCAAAGCATCTGTTTTTCTGAA
٦	128	168	184	180		7	189	229	245	241
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15

•				4	6/8	2				
acaatggaagctgttaaaggaaaacctgtttctgagcgggtgcaaagcatctgttatcgaa		TTGGCGCACCACTGTTATTAAGCTTAACGGTGTTTGCATTATTAATGATTTTTACGTCT	TTGGCGCACCTTATTAAGCTTAACGGTGTTTGCATTATTAATGATTTTAACGTCT	6 TTGGCGCAGCACTGTTATTAAGCTTAACGGTGTTTGCATTATTAATGATTTTTACGTCTTAATGATTTTTACGTCTTAATGATTTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTTTAATGATTAAT	TTGGCGCAGCACTGTTATTAAGCTTAACGGTGTTTGCATTATTTAATGATTTTTACGTCT	ttggcgcagcactgttattAAGCTTAACGGTGTTTGCATTATTTAATGATTTTTACGTCT	52 ATAATTTATATAGGATACAATCGATGAAAAACTTCTAATCGCAAGTTTATTATTCGGTAC	ATAATTTA	ATAATTTA	367 ATAATTTATAGGATACAATGGATGAAAAAACTTCTAATCGCAAGTTTATTTGGTGC
	7	250	290	306	302		53	311	351	36,
consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15

					47/8	2						
ATAATTTATATAGGATACAATGGATGAAAAAACTTCTAATCGCAAGTTTATTATTCGGTAC	ATAATTTATATAGGATACAATCGATGAAAAAACTTCTAATCGCAAGTTTATTATTCGGTAC	113 GACAACGACTGTGTTTGCCGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGGTGTTCAA	GACAACGACTGTGTTTGCCGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGTGTTCAA	GACAACGACTGTGTTTGCCGCACCTTTTGTGCCAAAAGATATTCGTGTGGATGGTGTTCAA	GACAACGACTGTGTTTGCCGCACCTTTTGTGCCAAAAGATATTCGTGTGGGATGGTGTTCAA		GACAACGACTGTGTTTGCCGCACCTTTTGTGGCAAAAGATATTCGTGTGGATGGTGTTCAA		1/4 GGTGACTTAGAACAACAAATCCGAGCAAGTTTACCTGTTCGTGCCGGTCAGCGTGTGACTG	GGTGACTTAGAACAACAAATCCGAGCAAGTTTACCTGTTCGTGCCGGTCAGCGTGTGACTG	473 GGTGACTTAGAACAAATCCGAGCAAGTTTACCTGTTCGTGCCGGTCAGCGTGACTG	
363		113	372	412	428	424			T / 1	433	473	
sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	L 7	cadio	minnad15	eagand15	

		•	Δ	8/82					•	
489 GGTGACTTAGAACAACAAATCCGAGCAAGTTTACCTGTTCGTGCtGGTCAGCGTGTGACTGACTGACTGACTGACTGACTGACTG	GGTGACTTAGAACAAAATCCGAGCAAGTTTACCTGTTCGTGCcGGTCAGCGTGTGACTG	ACAATGATGTGGCTAATATTGTCCGCTCTTTATTCGTAAGTGGTCGATTCGATGTGAA	ACAATGATGTGGCTAATATTGTCGGCTCTTTATTCGTAAGTGGTCGATTCGATGTGTAAA	ACAATGATGTGGCTAATATTGTCCGCTCTTTATTCGTAAGTGGTCGATTCGATGATGTGAA 🗭			ACAATGATGTGGCTAATATTGTCCGCTCTTTATTCGTAAGTGGTCGATTCGATGTGTGAA	AGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	555 AGCGCATCAAGAAGCGCGATGTGCTTGTTAGCGTTGTTGGCTAAATCGATCATTTCAGAT	
489		235	494	534	550	546		296	555	
pakd15 sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	

cad15

595 AGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	611 AGCGCATCAAGAAGGCGATGTTTTTTTTTTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	607 AGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	AGCGCATCAAGAAGGCGATGTGCTTGTTAGCGTTGTGGCTAAATCGATCATTTCAGAT	AAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAACTTAGATGCTA	616 GTTAAAATCAAAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAACTTAGATGCTA @	656 GTTAAAATCAAAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAACTTAGATGCTA	672 GTTAAAATCAAAGGTAACTCTGTTATTCCCACTGAAGCACTTAAACAAAACTTAGATGCTA	668 GTTAAAATCAAAGGTAACTCTATTATTCCACCTGAAGCACTAAAACAAAACTTAGATGCTA	GTTAAAATCAAAGGTAACTCTGTTATTCCcaCTGAAGCACTLAAACAAAACTTAGATGCTA	
eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus	

677 ACGGGTTTAAAGTTGGCGATGTTTTAATTCGAGAAAATTAAATGAATTTGCCAAAAGTGT	ACGGGTTTAAAGTTGGCGATGTTTTAATTCGAGAAAATTAAATGAATTTGCCAAAAGTGT	ACGGGTTTAAAGTTGCCGATGTTTTAATTCGAGAAAATTAAATGAATTTGCCAAAAGTGT		ACGGGTTTAAAGTTGGCGATGTTTTAATTCGAGAAAAATTAAATGAATTTGCCAAAAGTGT	0/8	479 AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACAGTTGAACCTATTGTCAATACGN	AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACAGTTGAACCTATTGTCAATACG	AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACAGTTGAACCTATTGTCAATACG	AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACCGTTGAACCTATTGTCAATACG	AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACCGTTGAACCTATTGTCAATACG	AAAAGAGCACTATGCAAGTGTAGGTCGCTATAACGCAACAGTTGAACCTATTGTCAATACG
677	717	733	729			479	738	778	794	790	
minnad15	eagand15	pakd15	sb33d15	consensus		cad15	minnad15	eagand15	pakd15	sb33d15	consensus

cad15	540	CTACCAAATAATCGCGCTGAAATTTTAATTCAAATGAAGATGATAAAGCAAAATTGG
minnad15	799	799 CTACCAAATAATCGCGCTGAAATTTTAATTCAAATCAATGAAGATGATAAAGCAAAATTGG
eagand15	839	CTACCAAATAATCGCGCTGAAATTTTAATTCAAATCAATGAAGATGATAAAGCAAAATTGG
pakd15	855	
sb33d15	851	
consensus		CTaccaaataatcgcgctgaaattttaattcaaatcaatgaagatgataaagcaaaattgg © N
cad15	601	CATCATTAA
minnad15	860	 \TTAA
eagand15	006	CATCATTAACTTTCAAGGGGAACGAATCTGTTAGTAGCAGTACATTACAAGAACAAATGGA
pakd15	916	CATCATTAA
sb33d15	912	CATCATTAACTTTCAAGGGGAACGAATCTGTTAGTAGCAGTACATTACAAGAACAAATGGA
consensus		CATCATTAACTTTCAAGGGGAACGAATCTGTTAGTAGCAGTACATTACAAGAACAAATGGA

662 ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATAAAT	921 ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATTATAAATTTGAAGGTGCGCAATTCGAG	961 ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATTAAAATTTGAAGGTGCGCAATTCGAG	977 ATTACAACCTGATTCTTGGTGGAAATTATGGGGGAAATTAAATTTGAAGGTGCGCAATTCGAG	973 ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATTAAAATTTGAAGGTGCGCAATTCGAG	U ATTACAACCTGATTCTTGGTGGAAATTATGGGGAAATATTTGAAGGTGCGCAATTCGAG N Q A	723 AAAGATTTGCAGTCAATTCGTGATTATTTTAAATAATGGCTATGCCAAAGCACAAATTA	982 AAAGATTTGCAGTCAATTCGTGATTATTTAAATAATGGCTATGCCAAAGCACAAATTA	1022 AAAGATTTGCAGTCAATTCGTGATTATTTTTAAATAATGGCTATGCCAAAGCACAAATTA	1038 AAAGATCTGCAGGCAATTCGTGATTATTTTTAAATAATAGCCTATGCCAAAGCACAAATCA	1034 AAAGATTTGCAGGCAATTCGTGATTATTTAAATAATGCCTATGCCAAAGCACAAATCA	AAAGATtTGCAGtCAATTCGTGATTATTTAAATAATGGCTATGCCAAAGCACAAATtA
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

784	CTAAAAC	1083		1095	CTAAAaCGGATGTTCAGCTAAATGATGAAAAAAAAAGTTAATGTAACCATTGATGTAAA		1104 TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT	1144 TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT	1160	1156	TGAAGGTTTACAGTATGACCTTCGTAGTGCACGCATTATAGGTAATCTGGGAGGTATGTCT
cad15	minnad15 1043	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15 1144	pakd15	sb33d15	consensus

cad15 eagand15 pakd15 sb33d15 consensus cad15 minnad15 eagand15		
sb33d15	1278	TTGCAGATGTAGAAAATGCAATTAAAGCAAAACTTGGGGAACGAGGTTACGGTAACACAAC
consensus		TTGCAGATGTAGAAAATGCAATTAAAGCAAAACTTGGAGAACGCGGTTACGGTAGCGAAC

GATGCTGGACGATTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG)) ;	consensus
GATGCTGGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG	1404	pakd15
GATGCTGGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTTGCTG	1388	eagand15
GACGACGTTTAACTGTTCGCCAACTTCGCT	1348	minnad15
GATGCTGGACGACGTTTAACTGTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTG	cad15 1089	cad15
ggtaaattcagtacctgattttgatgatgcaaataaaacattagcgataaccttgttt (y g		consensus
AGTAAATTCTGTACCTGATTTTGACGATGCAAATAAAACATTAGCGATAACCTTTGTTGTT	1339	sb33d1.5
AGTAAATTCTGTACCTGATTTTGACGATGCAAATAAAACATTAGCGATAACCTTTGTTGTT	1343	pakd15
GGTAAATTCAGTACCTGATTTTGATGATGCAAATAAAACATTAGCGATAACCTTGTTGTTGTTGTT	1327	eagand15
	1287	minnad15
단 -	1028	cadib

					56/82	>					
			S ATAGTACTTTACGTCAGGAAATGCGACAACAAGAAGGAACTTGGTATAATTCACAATTAGT	. •	ATAGCACTTTACGTCAGGAAATGCGCCAACAAGAAGGAACTTGGTATAATTCACAATTAGT W O	1 TGAGTTAGGAAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAAACAGTCGAAATT					TGAGTTAGGAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTcGAAAACCGAATT
1150	1409	1449	1465	1461		1211	1470	1510	1526	1522	
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

GATCCTATCAATGGTAGTAATGATGAAGTGGATGTCGTATATAAAGTCAAAGAACGTAACA	GATCCTATCAATGGTAGTGATGATGATGTCGTATATAAAGTCAAAGAACGTAACA	GATCCTATCAATGGTAGTGATGATGATGTCGTATATAAAGTCAAAGAACGTAACA	GATCCTATCAATGGTAGCAATGATGATGTCGTATATAAAGTCAAAGAACGTAACA	GATCCTATCAATGGTAGCAATGATGATGGATGTCGTATATAAAGTCAAAGAACGTAAACA	GATCCTATCAATGGTAG LAATGATGAAGTGGATGTCGTATATAAAGTCAAAGAACGTAAACA S	32	CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGTGGTATTAGTTATCAAGCAAG	CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGTGGTATTAGTTATCAAGCAAG	CGGGTAGTATCAACTTTGGTTACGGTACAGAGTGGTATTAGTTATCAAGCAAG	CGGGTAGTATCAACTTTGGTTTACGGTACAGAGAGTGGTATCAGTTATCAAACAAG	CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGAGTGGTATtAGTTATCAAGCAAG	CGGGTAGTATCAACTTTGGTATTGGTTACGGTACAGAGAGTGGTATtAGTTATCAAgCAAG
1272	1531	1571	1587	1583			1333	1592	1632	1648	1644	
cad15	minnad15	eagand15	pakd15	sb33d15	consensus		cad15 1333	minnad15	eagand15	pakd15	sb33d15	consensus

TGTTAAACAAGATAATTTCTTGGGAACAGGGCGGCAGTAAGTA		TGTTAAACAAGATAATTTCTTGGGAACAGGGCCGCAGTAAGTA	TATTAAACAAGATAATTTCTTGGGAACAGGGCGGCAGTAAGTA	TGTCAAACAAGATAATTTCTTGGGAACAGGGCGGCGGCAGTAAGTA	TgTtaaacaagataatttcttgggaacagggggggggggg	GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAA	GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAAA	GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAAA	GATTATGGTACGAGTGTCATTTGGGTTATACCGAaCCCTATTTTACTAAGATGGTGTAAA	GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAA	GATTATGGTACGAGTGTCAATTTGGGTTATACCGAGCCCTATTTTACTAAAGATGGTGTAA
1394	1653	1693	1709	1705		1455	1714	1754	1770	1766	
cad15 1394	minnad15	eagand15	pakd15 1709	sb33d15	consensus	cad15	minnad15	eagand15 1754	pakd15 1770	sb33d15 1766	consensus

1516 GTCTTGGTGGAAATGTTTTCTTTGAAAACTACGATAACTCTAAAAGTGATACATCCTCTAA	GTCTTGGTGGAAATGTTTTCTTTGAAAACTACGATAACTCTAAAAGTGATACATCCTCTAA	GTCTTGGTGGAAATGTTTTTTGAAACTACGATAACTCTAAAAGTGATACTCTTAA	GTCTTGGTGGAAATATTTTTTTTTTTTTTTTTTTTTTTT		GTCTTGGTGGAAATGTTTTTTTTTTTTTTTTTTTTTTTT	CTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAATAAC	CTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAATAAC	CTATAAGCGTACGACTTACGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAATAAC	CTATAAGCGTACGACTTATGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAAAAC	CTATAAGCGTACGACTTATGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAATAAC	CTATAAGCGTACGACTTAcGGAAGTAATGTTACTTTAGGTTTCCCTGTAAATGAAAATAAC
1516	1775	1815	1831	1827		1577	1836	1876	1892	1888	
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15	minnad15	eagand15	pakd15	sb33d15	consensus

	1760 TGATTTTCTTTTGGTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCAACTAAAGGG	FFCT	TGATTTTTCTTTTGGTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCAACTAAAGGG	TGATTTTTCTTTTGGTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCAACTAAAGGG	TGATTTTTCTTTTGGTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCAACTAAAGGG	TGATTTTTCTTTTGGAACTATAACAGCCTTAATAGAGGCTATTTCCCAACTAAAGGG 9	GTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGTTCTGATAACAAATACTACAAAC		GTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGTTCTGATAACAAATACTACAAAC			GTTAAAGCAAGTCTTGGTGGACGAGTTAC LATTCCAGGTTCTGATAACAAATACTACAAAC
	1760	2019	2059	2075	2071		1821	2080	2120	2136	2132	
•	cad15	minnad15 2	eagand15	pakd15	sb33d15	consensus	cad15 1821	minnad15	eagand15	pakd15	sb33d15	consensus
												•

TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCTCTGGGTTGTATCTGC	TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCTCTGGGTTGTATCTGC	TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCTCTGGGTTGTATCTGC	TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCGCTGGGTTGTATCTGC	TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCCCTGGGTTGTATCTGC	TAAGTGCAGATGTACAGGGTTTCTACCCATTAGACAGAGATCACCCCTGGGTTGTATCTGC N	AAAAGCATCTGCAGATATGCAAATGGTTTTGGAAACAAGCGTTTACCGTTCTATCAAACT	AAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAAGCGTTTACCGTTCTATCAAACT	AAAAGCATCTGCAGGATATGCTTTTTGGAAACAAGCGTTTACCGTTCTATCAAACT	AAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAAGCGTTTACCGTTCTATCAAACT	AAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAAGCGTTTACCGTTCTATCAAACT
1882	2141	2181	2197	2193		1943	2202	2242	2258	2254
cad15	minnad15	eagand15	pakd15	sb33d15	consensus	cad15 1943	minnad15	eagand15	pakd15	sb33d15

AAAAGCATCTGCAGGATATGCAAATGGTTTTGGAAACAAGCGTTTTACCGTTCTATCAAACT	cad15 2004 TATACAGCGGTGGCATCGGTTCATTACGTGGTTTTGCTTATGGTAGTATTGGACCTAACG	minnad15 2263 TATACAGCGGGTGGCATCGGTTCATTACGTGGTTTTTGCTTATGGTAGTATTGGACCTAACG	TATACAGCGGGTGGCATCGGTTCATTACGTGGTTTTGCTTATGGTAGTATTGGACCTAACG	pakd15 2319 TATACAGCGGGTGGCATTGGTTCATTACGCGGTTTTGCTTATGGTAGTATTGGGCCTAAtG	TATACAGCGGGTGGCATTGGTTCATTACGCGGTTTTGCTTATGGTAGGATTGGGCCTAACG	TATACAGCGGGTGGCAT c GGT r CATTACG t GGT r TTTGCT r ATGGTAG t AT r GGGCCTAA c G $\overset{ ext{N}}{\sim}$
•	2004	2263	2303	2319	2315	
consensus	cad15	minnad15	eagand15 2303	pakd15	sb33d15	consensus

THE COMPANDED AND CONTROLL OF THE TRANSPANDANCE OF THE SECOND OF THE SEC		TGCCGAATATGGTAGTGGTACTGGTACTTTTAAGAAGATAAGTTCTGA	TGCCGAATATGGTAATGGTAGTGGTACTGGTACTTTTAAGAAGATAAGTTCTGA		CTTTTAATAAGATAAGTTCTGA	
CAATTTATGCCGAATATGGTAATGGTAGTG		CAATTTATGCCGAATATGGTAATGGTAGTG	CAATTTATGCCGAATATGGTAATGGTAGTG		CAATTTATGCCGAACATGGTAATGGTA	
2065		2324	2364	 	2380	
Cad15) ; ; ;)	minnad15	eagand15	ง ก ง	pakd15) 1

sb33d15 2376 CAATTTATcaaGgtCAaaaTAAT

aaaTTTAATAAGATAAGTTCTGA

TGTGATTGGTGGTAATGCAATCGCLACAGCCAGCAGCTCATTTGTACCAACTCCATTTA

TGTGATTGGTGGTAATGCAATCGCLACAGCCCAGCCAGAGTTAATTGTGCCAACTCCATTTN **CAATTTATGccGaatAtggTAATggtagtggtactggtactTTTAAgAAGATAAGTTCTGA** TGTGATTGGTGGTAATGCAATCgCtACaGCtAGcGCAGAgtTaATTGTgCCAACTCCATTT TGTGATTGGTGGTAATGCAATCGCTACAGCTAGCGCAGAGTTAATTGTGCCAACTCCATTT TGTGATTGGTGGTAATGCAATCGCTACAGCTAGCGCAGAGTTAATTGTGCCAACTCCATTT TGTGATTGGTGGTAATGCAATCGCTACAGCTAGCGCAGAGTTAATTGTGCCAACTCCATT 2126 2429 2422 2385 2425 cad15 pakd15 sb33d15 minnad15 eagand15 consensus consensus

GTGAGCGATAAGAGCCAAAATACGGTCCGAACCTCCTTATTTGTTGATGCGGCAAGTGTTT GTGAGCGATAAGAGCCAAAATACGGTCCGAACCIICCTTATTTGTTGATGCGGCAAGTG'I'T'I CTCCTTATTTGTTGATGCGGCAAGTGTTT GTGAGCGATAAGAGCCAAAATACGGTCCGAACC 2187 2446 2486 cad15 eagand15 minnad15

2568

minnad15

pakd15 2490 GTGAGTGATAAAAGCCAAAATACAGTCCGAACCTCCCTATTTGTTGATGCGGCAAGTGTTT 	GTGAGcGATAAgAGcCAAAATACgGTCCGAACCTCCtTATTTGTTGATGCGGCAAGTGTTT	GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCGATGTATTAAAAAGATTGCC		3AATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCGATGTATTAAAAAGATTGCC N	GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCAAGGTCTTGAAAGACTTACC	GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCAAtGTCTTGAAAGACTTACC	GGAATACTAAATGGAAATCAGATAAAAATGGATTAGAGAGCgAtGTaTTaAAAAagaTTgCC	cad15 2309 TGATTATGGCAAATCAAGCCGTATTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCT
GTGAGTC GTGAGTG	GTGAGcC	GGAATAC	GGAATAC			GGAATAC	GGAATAC	TGATTAT
2490		2248			2551	2544		2309
pakd15 sb33d15	consensus	cad15	minnad15 2507	eagand15 2547	pakd15	sb33d15	consensus	cad15

2431

cad15

eagand15	2608	eagand15 2608 TGATTATGGCAAATCAAGCCGTATTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCT
pakd15	2612	TGATTATGGC
sb33d15	2605	
consensus		tGATTATGGCAAATCAAGCCGTAtTCGCGCCTCTACAGGTGTCGGATTCCAATGGCAATCT
cad15	cad15 2370	CCTATTGGGCCATTGGTATTCTCTTTATGCCAAACCAATTAAAAAATATGAAAATGATGATGOTGOTGOTGOTGOTGOTGOTGOTGOTGOTGOTGOTGOT
minnad15	2629	CCTATTGGGCCATTGGTATTCTCTTTATGCCAAACCAATTAAAAAATATGAAATGATGATGATG IIIIIIIIIII
eagand1.5	2669	CCTATTGGGCCATTGGTATTCTCTTTATGCCAAACCAATTAAAAAATATGAAAATGATGATG
pakd15 2673	2673	
sb33d15	2666	
consensus		CCTALTGGGCCALTGGTATTcTCTTATGCcAAACCAATTAAAAAATATGAAAATGATGATG

				6	57/8	32					
0 TCGAACAGTTCCAATTTAGTATTGGAGGTTCTTTCTAATAAATTGAACTTTTTTCTTCAT	0 TCGAACAGTTCCAATTTAGTATTGGAGGTTCTTTCTAATAAATTGAACTTTTTTTT	pakd15 2734 TCGAACATTCCAATTTAGTATTGGGGGCTCTTTCTAATAAATTGAACTTTTTCGTCATC	7 TCGAACAGTTCCAATTTAGTATTGGGGGTTCTTTCTAATAAATTGAACTTTTTTCGTCATC	TCGAACAGTTCCAATTTAGTATTGGAGGLTCTTTCTAATAAATTGAACTTTTTTCLTCATC		2 AGAACTCAAAAACAACGTTCTCTGCCTAATTTAATTGGGCAGAGAAATATTAAACCCAT	1 AGAACTCAAAAACGTTCTCTCTGCCTAATTTAATTGGGCAGAGAAATATTAAACCCAT	1 AGAACTCAAAAACAACGTTCTCTGCCTAATTTAATTGGGCAGAGAAATTAAATTAGGCAGAGAAATTTAAACCCATG	akd15 2795 AGAACTCAAAAACGACGTTCTCTCTGCCTAATTGAATTG		AGAACTCAAAAACAACGTTCTCTGCCTAATTtAATTGGGCAGAGAAAATATTAAAcCCATC
2690	2730	273	2727		,	677	2751	2791	2795	2788	
minnad15	eagand15	pakd15	sb33d15	consensus	, , ,	Cadio	minnad15	eagand15	pakd15 2795	sb33d15	consensus

cad15 2553 minnad15 2812 eagand15 2852 pakd15 2856 sb33d15 2849 consensus cad15 2614 minnad15 2873 eagand15 2913 pakd15 2917 sb33d15 2910	2553 2812 2852 2856 2849 2614 2873 2913 2917	ATTTAATTAAGGATATTTATCAAATGAAAAACATCGCAAAAGTAACCGCACTTGCTTTAGG ATTTAATTAAGGATATTTATCAAATGAAAAAAACATCGCAAAAGTAACCGCACTTGCTTTAGG ATTTAATTAAGGATATTTATCAAATGAAAAAAAAAGTAACCGCACTTGCTTTAGG ATTTAATTAAGGATATTTATCAAATGAAAAAAAAAGTAACCGCACTTGCTTTAGG ATTTAATTAAGGATATTTATCAAATGAAAAAAAAAGTAACCGCACTTGCTTTAGG ATTTAATTAAGGATATTTATCAAATGAAAAAAAAAGCAAAGTAACCGCACTTGCTTTAGG ATTTAATTAAATTA
consensus		TaTTGCACTTGCTTCAGGCTATGCTtCcGCTGAAGAAAAAATTGCTTTcATTAATGC-agt

თ _---gatcgccaagcggtagcagataaacttgatgctgaatttaaდ N atattTTTcaAcatCacccagatcgccaagcggtagcagataaacttgatgctgaatttaa TATAnTTTnCAAggCnaagg ttat-ttttcaaa-c--TATALTTTCAA TTATA cad15 2675 2934 2973 2970 pakd15 2977 minnad15 eagand15 sb33d15 consensus

cad15 2736 acctgtagctgagaaattagcagcaagcaaaaaagaagttgatgataaaattgctgctgct

minnad15 2954

eagand15 2985

pakd15 2990

sb33d15 2975

consensus

acctgtagctgagaaattagcagcaagcaaaaaaaagaagttgatgataaaattgctgctgct

FIG.1F.(CONTINUED)

cad15 2797 cgtaaaaaagtagaagcaaaagttgcggctttagaaaaagatgcacctcgcttacgtcaag

minnad15 2954

eagand15 2985

pakd15 2990

sb33d15 2975

consensus

ctgatattcaaaaacgccaacaggagattaataaattaggtgcggctgaagatgctgaatt 2858 cad15

minnad15 2954

eagand15 2985

pakd15 2990

sb33d15 2975

ctgatattcaaaaacgccaacaggagattaataaattaggtgcggctgaagatgctgaatt consensus

71/82

FIG.1F.(CONTINUED)

cad15 2919 acaaaaattaatgcaagaacaagataaaaa

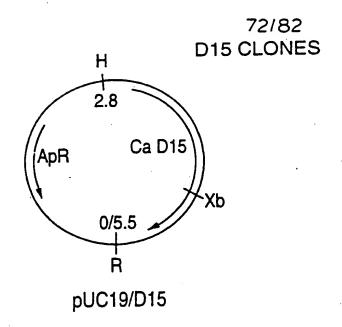
minnad15 2954 eagand15 2985

pakd15 2990

sb33d15 2975

acaaaaattaatgcaagaacaagataaaaa consensus

SUBSTITUTE SHEET



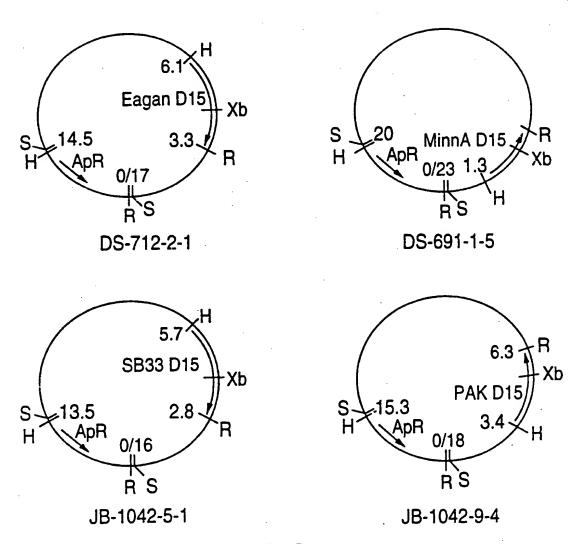


FIG.2.
SUBSTITUTE SHEET

D15 SEQUENCE COMPARISON

	_	_				_	~		73/	82	_	4					~		
පු	Eagan	MinnA	SB33	PAK	2	Eagan	MirmA	SB33	PAK	යි	Eagan	MinnA	SB33	PAK	යි	Eagan	MinnA	SB33	PAK
MRTILIASII.FGTTTTVFAAPFVAKDIRVDGVQGDI.FQQIRASI.FVRAQQRVTIDNDVANIVRSI.FVSGREDDVKAHQEGDVLVVSVVAKSII.SDVKIKGN					SVIPTEALKONLDANGERKGEDVLIREKINEFAKSVKEHYASVERYNATVEPTANII.PANRAETLIOINEDDKAKI ASIJIFKRANESVSSSII OEDMEI OED			I.		SMWKIMENKEEGAQFEKDIQSIRDYYIMNGYAKAQITKIDXQINDEKTKANVITIDANEGIQYDIRSARIIGNIGGAKSAELEPLISALHINDIFRRSDIAD			**************************************		VENATKAKLGERGYGSATVNSVPDFTDANKTI ALTITAVDAGRRUTVRQI RFEGNIVSADSTI RQEMRQOEGIWYNSOLVELGKI RLDRIGFFETVENRID				FFF

FIG.3/

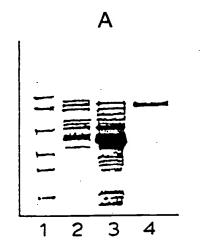
SB33 PAK	**
Eagan	**
ප	PTPFVSDKSQNTVRISLFVDAASVANTKAKSDKN3LESDVLKRLPDYCEKSSRLRASTGNGFQAQSPIGPLVFSYAKPIKKYENDDVEQFQFSIGGSF**
PAK	R
SB33	
MinnA	
82 ප	YKLSADVQGFYPLDRDHLMVVSAKASAGYANGFGNKRLPFYQTYTAGGIGSLRGFAYGSIGRNALYAEYGNGSGIGTFKKLSSDVIGGNALATASAELIV
74 / Yea	
SB33	
MirmA	
Eagan	
පී	VYKRTTYGSNVTLGFPANENNSYYVGLGHTYNKLISNFALEYARNLYTQSMKFKGNGIRTNDFDFSFGMAYNSIARGYFPTKGVKASLGGRVTTFGSDNKY
PAK	T.T.
SB33	
MirmA	
Eagan	
පු	PINGSNDEVDVVYKVKERNIGSINFGIGYGTESGISYQASVKQINFLGIGAAVSIAGIKNDYGISVNLGYTEPYFTKIGVSLGQAVFFENYINSKSDTSS

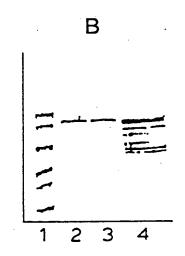
FIG.3F

75/82 Construction of plasmid expressing SB33 D15 ApR pUC JB-1042-5-1 13.5 2.8 ApR 0/16 R/H RH H BsrF I ApR D15 'ApR pUC BsrF I pRY-60-1 BsrF I Nde RH BsrF I/R R/Nde Nde-BsrF I oligos Nde ∠BsrF I ApR **ApR** ApR D15 pUC BsrF I DS-860-1-1 pT7-7 -Xb \\rangle R Nde Вġ Bg/Nde R/Rg Nde/R H/ Nde D15 ApR DS-880-1-2 Xb FIG.4.

SUBSTITUTE SHEET

76/82
PURIFICATION OF D15 FROM A NON-TYPEABLE HAEMOPHILUS INFLUENZAE STRAIN 30





PROTEIN STAIN

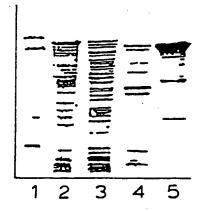
WESTERN BLOT

- 1. Low MW markers
- 2. Strain 30
- 3. Native D15 crude extract
- 4. D15 after anti-D15 affinity chromatography

FIG.5.

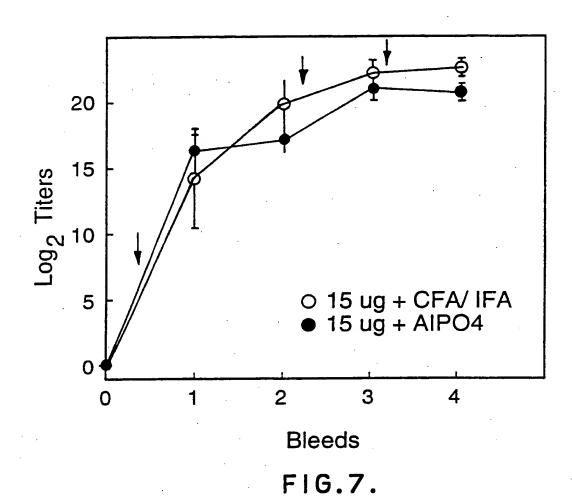
77/82

PURIFICATION OF FULL LENGTH RECOMBINANT D15



- 1. Protein M.W. Markers
- 2. Lysate of E. coli expressed rD15
- 3. Soluble protein in Tris-HC1 buffer extract
- 4. Soluble proteins in Tris/Triton X-100/ EDTA extraction buffer
- 5. rD15 inclusion bodies

FIG.6.



79/82

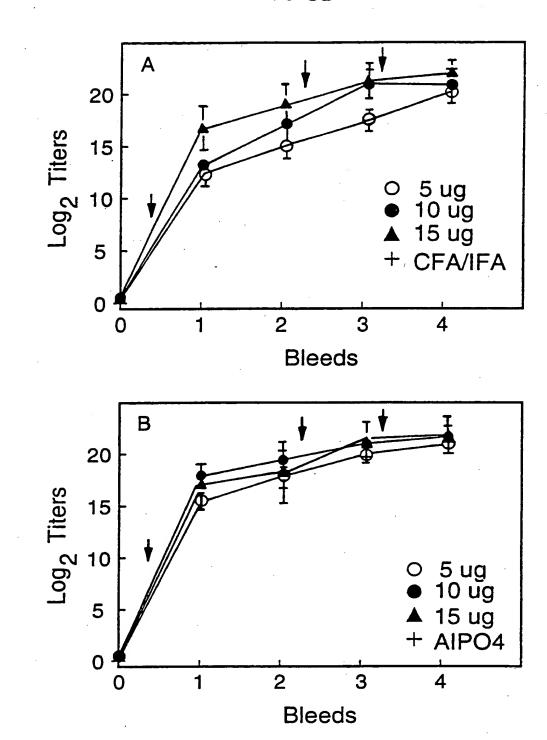
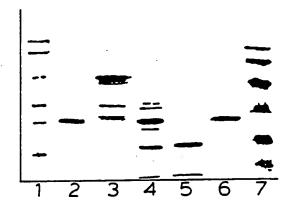


FIG.8.

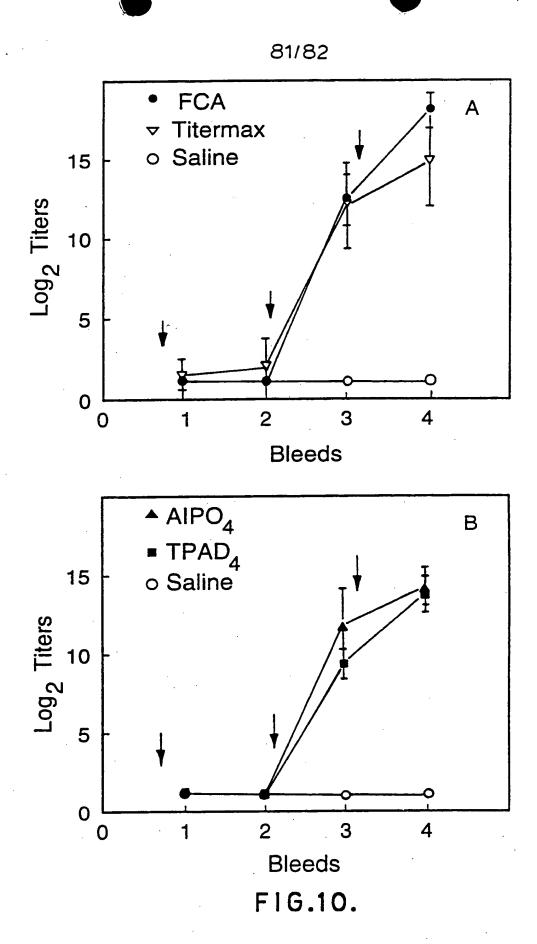
80/82

PURIFICATION OF TRUNCATED D15 FROM D15-GST FUSION PROTEIN



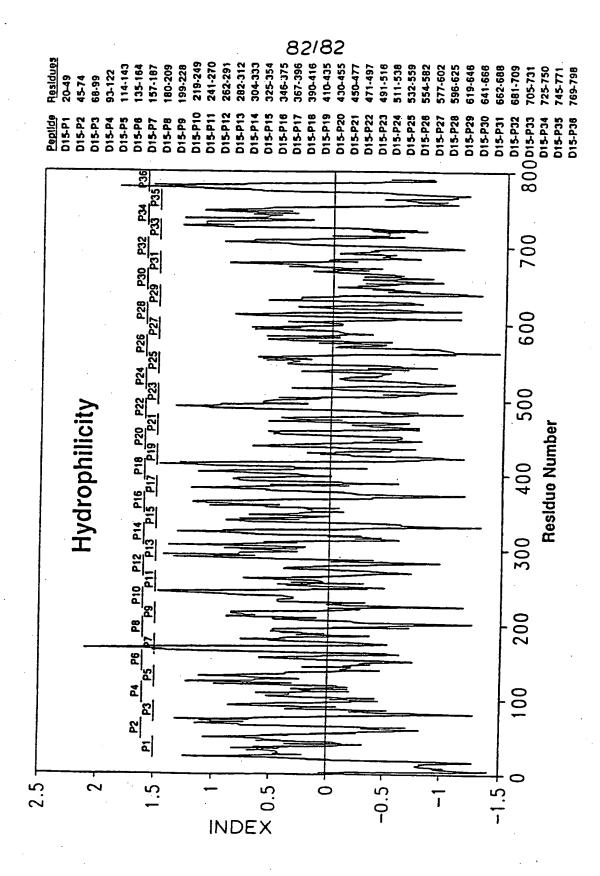
- 1. Prestain low MW markers
- 2. GST standard
- 3. GST-(D15 fragment) fusion protein
- 4. Fusion protein cleaved by thrombin
- 5. rD15 fragment
- 6. GST
- 7. Low MW markers

FIG.9.



SUBSTITUTE SHEET





INTERNATIONAL SEARCH REPORT



ional Application No

		•		PCT/CA	93/00501
at the state of th	SIFICATION OF SUBJECT MATTER				
C :	12 N 15/31,C 07 K 13/00,A 61 (C 12 N 15/31; C 12 R 1/21)	K 39/102	•		
1	to International Patent Classification (IPC) or to both national	classification and IPs	C		
	S SEARCHED	The state of the s			
Minimum	documentation searched (classification system followed by class	ification symbols)			
A	51 K.C 07 K.C 12 N				
Document	ation searched other than minimum documentation to the extent	that such documents	are included in the field	ts searched	
	•	•			
Electronic	data base consulted during the international search (name of dat	a base and, where pr	actical, search terms us	×0	
		•	•	•	•
C. DOCUN	MENTS CONSIDERED TO BE RELEVANT				
Category *	Citation of document, with indication, where appropriate, of t	he relevant passages		Releva	at to claim No.
A	THERCHTON AND THRONIS				•
A	INFECTION AND IMMUNITY, vol. 58, no. 4, Apri	1 1000		1,4	,6,
*	W.R. THOMAS et al. "	EVDres-		27	
	sion in Escherichia	coli of a	•		
	High-Molecular-Weigh	it	•	·	
	Protective Surface A	ntigen			
	Found in Nontypeable	and		1	
	Type b Haemophilus i enzae"	nriu-			
	pages 1909-1913,				
	the whole document.	•			
А	ED 32 0 270 000				
^	EP, A2, 0 378 929 (CONNAUGHT LABORATOR	TRO		1,4,	
j	LIMITED) 25 July 199	U TES		12.1	
l	(25.07.90),	•		20.2	. Z
ĺ	claims.				
A	WO 31 01/06 655				
^	WO, A1, 91/06 652			1,4,	6,
				<u> </u>	· · · · · · · · · · · · · · · · · · ·
Furth	er documents are listed in the continuation of box C.	Patent fr	amily membas are listed	1 in annex.	
* Special cat	egones of ated documents:	app lates de sur			
A' docume	nt defining the general state of the art which is not	or priority d	ent published after the in late and not in conflict v	with the applican	on but
COUNTE	red to be of particular relevance locument but published on or after the international	invention	erstand the principle or	•	Ĭ
ming o	ate nt which may throw doubts on priority claim(s) or	Cannot be ex	particular relevance; the considered power or cannot be considered.	ot be considered	1 0
ALECT D	s cited to establish the publication date of another or other special reason (as specified)		nventive step when the o particular relevance; th		
O, qocume	nt referring to an oral disclosure, use, exhibition or	cannot be or	ensidered to unvolve an i	aveatre step wh	en the . I
oeser m	cans nt published prior to the international filing date but	ments, such	combination being obve	ous to a person s	trilled
OLEF (B)	in the priority date claimed		ember of the same pares	it femily	
Date of the a	ctual completion of the international search	Date of mail	ing of the international s	earth report	
	15 February 1994		25 -03- 1994		İ
			1334		
Name and mi	siling address of the ISA	Authorized of	flicer		
	European Patent Office, P.B. 5818 Patentiaan 2 NL - 2280 HV Rijswijk				į
	Tel. (+31-70) 340-2040, Tx. 31 651 epo nl. Fax: (+31-70) 340-3016	1	WOLF e.h.		i

ANHANG

zum internationalen Recherchenbericht über die internationale Patentangeldung Nr.



ANNEX

to the International Search Report to the International Patent Application No.



NNEXE

au rapport de recherche international relatif à la demande de brevet international no

PCT/CA 93/00501 SAE 82214

In diesem Anhang sind die Mitglieder der Patentfamilien der im obengemannten internationalen Recherchembericht cited in the above-mentioned interangeführten Patentdokumente angegeben. Diese Angaben dienen nur zur Unterrichtung und erfolgen shoe Gewähr.

This Annex lists the patent family aembers relating to the patent documents national search report. The Office is in no way liable for these particulars which are given merely for the purpose of information.

La présente annexe indique les sempres de la famille de brevets relatifs aux documents de brevets cités dans le rapport de recherche international visée ci-dessus. Les reseignements fournis sont donnés à titre indicatif et n'engagent pas la responsibilité de l'Office.

ang: Do	eführte Patent im sea Cument	herchenbericht es Patentdokunent document cited urch report de brevet cité upport de recherche	Datum der Veröffentlichung Publication date Date de publication	ratent family aesber(s) Hesbre(s) de la	Datum der röffentlichung Publication date Date de publication
EP	A2	378929	25-07-90	CA AA 2006587 2 CA AA 2006587 2 EP A3 378929 0 JP A2 2283286 2	2-03-89 2-06-90 3-06-90 1-08-90 0-11-90 2-02-89
WD	A1	9106652	16-05-91	EP A1 500576 0 GB A0 8924473 20	1-05-91 2-09-92 0-12-89 1-03-93
EP	A1	281673	14-09-88	DK A0 6020/87 19	2-12-92 7-11-87 7-05-88 0-06-89
us 	A	5013664	07-05-91	keine – none – rien	